AD-A239 937

AGARD-AG-330

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARDograph 330

Considerations for NATO Satellite Communications in the Post 2000 Era

(Eléments de Réfléxion pour les Communications par Satellite dans les Pays Membres de l'OTAN après l'An 2000)

NORTH ATLANTIC TREATY ORGANIZATION

Distribution and Availability on Back Cover

AGARD-AG-330

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARDograph 330

Considerations for NATO Satellite Communications in the Post 2000 Era

(Eléments de Réflexion pour les Communications par Satellite dans les Pays Membres de l'OTAN après l'An 2000)

Edited by

Prof. Dr A. Nejat Ince
MARMARA Scientific and Industrial Research Center
P.O. Box 21
41470 Gebze-Kocaeli
Turkey

A-1

The material in this publication was assembled to report the results of the AGARD Avionics Panel Working Group 13.

North Atlantic Treaty Organization
Organisation du Traité de l'Atlantique Nord

The Mission of AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

— Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;

— Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);

— Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;

— Improving the co-operation among member nations in aerospace research and development;

— Exchange of scientific and technical information;

— Providing assistance to member nations for the purpose of increasing their scientific and technical potential;

— Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced
directly from material supplied by AGARD or the authors.

Published June 1991

ISBN 92-835-0622-7

*Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ*

Preface

Working Groups may be established if AGARD Panels feel that expertise which does not exist within its membership is required to address specific problems or to produce state-of-the-art reports of interest to NATO. Such a need arose in the Avionics Panel when the writer and Dr P. Bakken submitted a proposal in 1986 for a futuristic study on satellite communications.

Among the space activities of the last three decades, SATCOM has found the widest application in meeting both civil and military communications requirements. Several international, regional and national SATCOM systems of increasing capacity, capability and complexity have been and are being implemented over the years. The latest versions are utilizing such concepts as spot beams, processing transponders in SS-TDMA and operations in different frequency bands including the EHF band.

On the military side, the United States, the United Kingdom, France, and NATO have been the only owners and operators of military SATCOM systems in the West. The systems in being and under development use satellites and ground terminals with characteristics which differ from the civilian ones with respect to frequency bands utilized and survivability and interoperability. The SATCOM has given the military users the potential of having much-needed mobility, flexibility and survivability in strategic and tactical communications for land, sea and air operations. It must, however, be said particularly for the military SATCOM systems that they have been evolved in big jumps, both in time and capability, each jump involving the deployment of two or three often specially designed large satellites, large expenses and rather traumatic transition between jumps. Despite these undesirable features these systems did not have the required degree of survivability, and flexibility. Clearly an entirely new approach to system architectural design and acquisition is needed with emphasis on versatility, flexibility, resistance to electronic and physical attacks, cost-effectiveness and for the possibility of increased participation of NATO Nations in future cooperative R&D in SATCOM. This will also have to take into account the following facts and trends:

i) The use of SATCOM is expected to be more pervasive, particularly for small mobile users (aircraft, land mobiles, ships, and submarines) to support general purpose and modern C3I structures.

ii) There will be more data traffic.

iii) There will be a requirement for cheaper SATCOM.

iv) Frequencies in the EHF and optical bands will be needed for greater capability (e.g., communications with submerged submarines) and smaller terminals.

v) SATCOM will be integrated with the future all digital ISDN and this may require SATCOM to have improved effective performance (e.g., minimum delay, echo).

NATO has invested substantial sums in satellite communications and has therefore continuous interest in the subject particularly as far as future prospects are concerned. It is believed that the rapid and much welcome political changes that are taking place in Europe and elsewhere would not diminish the importance of SATCOM for NATO; on the contrary a SATCOM system with attributes outlined above would be a great asset for mobile forces and relaying of intelligence data in this less-tensioned but more uncertain era into which NATO is entering. One of the consequences of the recent political events is that planning for the evolution of NATO communications, which had previously been regarded as firm, is now likely to come under review.

These are then the justifications for NATO interest in an imaginative, basically technology-driven study of concepts for a future NATO SATCOM system which is mindful of what exists today but otherwise unconstrained by the past. It was recognised by the Avionics Panel and others in NATO that there are several bodies in NATO such as NACISA, SCSG (Satellite Communications Sub-Group), TSGCEE (Tri-Service Group on Communications and Electronic Equipment) and STC (SHAPE Technical Centre) which are involved in SATCOM work which, however, relate basically to the problems of existing SATCOM and its immediate future planning. In the light of all this information and in full consultation with the above bodies, the National Delegates Board (NDB) of AGARD approved in March 1986 the proposal by the Avionics Panel to establish an AGARD Working Group (WG-13) to consider the future potential of satellite communications in the NATO context.

The WG-13 thus created had the Terms of Reference given on page xi. The scientists and engineers who participated in the deliberations of the Group came from the R&D establishments and space companies of Canada, France, Germany, Norway, Turkey, the United Kingdom, the United States and from SHAPE Technical Centre and the International Military Staff. The STC representative undertook the task of informing the NATO bodies concerned with the activities of WG-13 as they occurred. The meetings of the Group were generally held at the establishments of the members. On these occasions opportunities were always taken to visit laboratories, receive briefings and collect information together related to satellite communications.

The report that follows is the joint work of all the members of the Group and is believed to contain a substantial amount of information, some of it original, on almost every aspect of satellite communications (system architecture, design, devices, techniques and technologies for space, ground and system control segments) which, it is hoped, will be found useful by a wide range of readers including military and political decision makers, planners, and the R&D community.

The SATCOM system concepts considered have the attributes required for both strategic and tactical communications They are flexible, modular and possess the capability to change with changing requirements in an evolutionary manner using today's technology but without expensive developments.

For completeness' sake one must mention the idea, being considered today, of using small satellites (300—600 kg, 2—12 month lifetime, 4—12 hr elliptical orbits) either as a complement to larger systems or stand-alone crisis systems for theatre operations. This concept is also embraced by the architectures and technologies, for both spacecraft and launch vehicles, described in this report.

The work took almost three years to complete and demanded a lot of studies, discussions, evaluation of options, development of new ideas and concepts and writing reports from a group of exceptionally qualified but very busy people who had to find, and sometimes create time, and to devote effort beyond the call of their normal duties, just to serve the cause of NATO.

As the director privileged to work with such a devoted and talented group I would like to record here my deep personal appreciation and gratitude to the members of WG-13 and the NATO Nations who so generously supported them by, in the first place, nominating them to the Group and then providing information on national R&D work and hosting the meetings of the Group. Last but by no means least, on behalf of the Group I would like to thank sincerely AGARD and the Avionics Panel and its Executive Officer for all the encouragement and financial and administrative support provided.

I hope that the reader will find this report as useful as the Group has found it enjoyable and worthwhile to produce.

Prof. Dr Nejat Ince
Director
Working Group 13

Préface

Des groupes de travail peuvent être créés lorsque les Panels de l'AGARD estiment que la compétence requise pour l'examen de certains problèmes particuliers ou pour la rédaction de rapports sur l'état de l'art dans tel ou tel domaine n'existe pas au sein du Panel. Un tel cas s'est présenté pour le Panel d'Avionique en 1986, quand l'auteur et le Docteur P. Bakken ont soumis une proposition relative à une étude futuriste sur les télécommunications par satellite.

Parmi les différentes activités spatiales des trois dernières décennies, les SATCOM ont trouvé les plus larges applications dans le domaine des télécommunications civiles et militaires. Plusieurs systèmes régionaux, nationaux et internationaux SATCOM d'une puissance, d'une capacité et d'une complexité toujours grandissante ont été mis en service pendant cette période. Les dernières versions intègrent des concepts tels que les faisceaux ponctuels, les répéteurs non-transparents en SS-TDMA et l'exploitation en différentes bandes de fréquences y compris les ondes millimétriques.

Sur le plan militaire, les Etats-Unis, le Royaume-Uni, la France, et l'OTAN sont les seuls détenteurs et exploitants de systèmes SATCOM militaires occidentaux. Les systèmes en service ou en développement utilisent des satellites et des stations au sol dont les caractéristiques diffèrent des installations civiles en ce qui concerne les bandes de fréquences utilisées, la survivabilité et l'interopérabilité.

Les SATCOM offrent à l'utilisateur militaire des caractéristiques très demandées qui sont: la mobilité, la flexibilité et la survivabilité pour les télécommunicationsstratégiques et tactiques dans les opérations terre, mer et air.

Pourtant, il est à noter que, dans le cas des systèmes SATCOM militaires en particulier, l'évolution est caractérisée par des sauts temporels et qualitatifs de grande envergure. Chaque saut implique le déploiement de deux ou trois grands satellites de conception spéciale, l'engagement de capitaux considérables, et des périodes de transition plutôt stressantes entre les différentes phases des opérations.

Malgré ces facteurs indésirables, ces systèmes n'atteignent pas le niveau requis de survivabilité et de flexibilité. Il est clair qu'une nouvelle approche de la conception architecturale et de l'acquisition des systèmes est nécessaire, l'accent étant mis sur l'adaptabilité, la flexibilité, la résistance à l'intrusion électronique, le durcissement et la rentabilité en s'appuyant sur les possibilités d'une participation accrue des pays membres de l'OTAN aux activités futures de R&D dans le domaine des SATCOM. Cette approche devra tenir compte des contraintes et des tendances suivantes:

i) Utilisation plus généralisée des SATCOM, en particulier pour les usagers individuels mobiles, (aéronefs et mobiles terrestres, navires et sous-marins) en tant que soutien aux structures polyvalentes et C3I.

ii) Accroissement du trafic de données.

iii) Diminution des coûts des SATCOM.

iv) Utilisation d'ondes millimétriques et de bandes de fréquences optiques pour l'augmentation de la capacité (transmissions avec des sous-marins en plongée) et réalisation de centres terminaux plus petits.

v) Les SATCOM doivent être intégrés à tous les futurs réseaux numériques à intégration de services (ISDN) ce qui nécessitera sans doute des SATCOM plus performants (temps d'exécution minimal, echo etc...)

L'OTAN a investi des sommes importantes dans les télécommunications par satellite et continue donc à s'intéresser à ce sujet, en particulier pour l'avenir. Il semblerait que les heureux changements politiques qui s'effectuent en ce moment en Europe et ailleurs ne réduiront pas l'importance des SATCOM pour l'OTAN. Bien au contraire, un système SATCOM ayant les caractéristiques mentionnées ci-dessus constituerait un atout majeur pour les forces mobiles et pour la retransmission de données du renseignement à l'aube d'une ère moins critique pour l'OTAN, mais plus incertaine aussi.

L'une des conséquences des évènements politiques récents est la remise en question possible des plans concernant l'évolution des télécommunications OTAN considérés jusqu'ici comme définitifs.

Voici donc la justification de l'intérêt que pourrait montrer l'OTAN pour un futur système SATCOM OTAN qui tiendrait compte de la situation actuelle, mais qui ne serait pas obéré par le passé. Il a été admis par le Panel d'Avionique et par d'autres services de l'OTAN qu'il existe plusieurs organismes de l'OTAN, tels que NACISA SCSG (sous-groupe pour les télécommunications par satellite), TSGCEE (groupe tri-service pour les télécommunications et le matériel électronique) et STC (centre technique du SHAPE) qui travaillent sur les SATCOM (essentiellement en ce qui concerne les problèmes des SATCOM existantes et de leur avenir immédiat.

A la lumière de toutes ces informations et en consultation avec les organismes mentionnés ci-dessus, le Conseil des Délégués Nationaux de l'AGARD (NDB) a approuvé, au mois de mars 1986, la proposition soumise par le Panel d'Avionique de l'AGARD, visant à la création d'un groupe de travail (WG 13) qui réfléchirait au potentiel des télécommunications par satellite dans le cadre des activités de l'OTAN.

Le WG 13 ainsi créé a eu pour mandat les termes de référence contenus dans le présent document. Les scientifiques et ingénieurs qui ont participé aux délibérations du groupe venaient des établissements de recherche et développement et des sociétés aérospatiales du Canada, de la France, de l'Allemagne, de la Norvège, de la Turquie, du Royaume-Uni, des Etats-Unis, du Centre Technique du SHAPE et de l'Etat-Major International de l'OTAN. Le représentant du STC s'est chargé d'informer les organismes de l'OTAN concernés des progrès du groupe au fur et à mesure de leur avancement. Les réunions du groupe se sont tenues en général dans les établissements principaux des différents membres, et les membres du groupe profitaient de ces occasions pour visiter les laboratoires, pour assister à des exposés et, en général pour recueillir des informations concernant les télécommunications par satellite.

Le rapport qui suit représente un travail commun fourni par tous les membres du groupe. Il contient de très nombreuses informations, dont certaines sont inédites, sur pratiquement tous les aspects des télécommunications par satellite (architecture de système, conception, mécanismes, techniques et technologies pour le contrôle sectoriel spatial, terrestre et système). Il est à espérer que ce rapport répondra aux besoins d'un large éventail de lecteurs, dont les décideurs militaires et politiques, les planificateurs et la communaute R&D en général.

Les concepts de système SATCOM considérés ont les caractéristiques demandées pour les télécommunications stratégiques et tactiques. Ils sont flexibles, modulaires et ils ont la capacité d'évoluer selon les besoins du moment de la communauté R&D, en utilisant des technologies actuelles sans avoir recours à des développements coûteux.

Par souci d'exhaustivité, je citerais la notion, actuellement à l'étude, d'utiliser plusieurs petits satellites (300—600 kg, cycle de vie 2—12 mois, orbites elliptiques 4—12 h) soit en complément aux systèmes plus volumineux, soit en système autonome de crise, soit pour des théâtres d'opérations. Ce concept fait partie intégrante des architectures et technologies des vecteurs spatiaux et des lanceurs décrits dans ce rapport.

Presque trois ans ont été consacrés aux travaux, trois ans d'études, de discussions, d'évaluation d'options, de développement d'idées et de concepts nouveaux, et à la rédaction de rapports, de la part d'un groupe de personnes éminemment competentes mais très occupés, qui ont su trouver, sinon créer du temps et consacrer leurs efforts en plus de leurs charges de travail habituelles, tout simplement pour servir la cause de l'OTAN.

En tant que directeur du WG 13, je m'estime privilégié d'avoir travaillé en compagnie d'un groupe si dévoué et si talentueux. Je tiens à souligner ici ma profonde reconnaissance personnelle et mes remerciements aux membres du WG 13 et aux pays de l'OTAN qui les ont soutenu si efficacement, en les désignant d'abord comme membres du groupe, en les fournissant ensuite des informations concernant les activités R&D en cours dans les pays concernés, et en acceptant de prendre en charge l'organisation des différentes réunions. En dernier lieu, et de la part du groupe, je tiens à remercier très sincèrement l'AGARD, le Panel d'Avionique et son administrateur pour l'encouragement et le soutien financier et administratif qu'ils ont bien voulu nous fournir.

J'espère que le lecteur trouvera dans ce rapport autant d'utilité que les membres du groupe ont trouvé de plaisir et d'intérêt dans sa réalisation.

Prof. Dr Nejat Ince
Directeur
Groupe de Travail 13

Members of Working Group 13

Overall 14 highly qualified individuals nominated by 7 NATO countries, NATO HQ and SHAPE Technical Centre participated in the deliberations of WG 13. The complete list of participants is given below:

Prof. Dr A. Nejat Ince, Director (AVP)	Turkey
Dr Ottfried Baetz	Germany
Dr Petter M. Bakken (AVP)	Norway
Mr Jacques Chaumeron	France
Dr Barry Felstead	Canada
Dr Franklin W Floyd	United States
Mr F. William Jackson (AVP)	United Kingdom
Dr Dharan P.S. Malik	United Kingdom
Dr Paul Masterman (STC)	United Kingdom
Mr Jean-Louis de Montlivaut	France
Lt Col. Paul J. Rainville (IMS/NATO)	United States
Mr Rolf Rosenberg	Germany
Dr James Saunders	Germany
Dr Robert Schweikert	Germany

Contents

Terms of Reference

for AVP Working Group 13 on Satellite Communications

BACKGROUND

At their Business Meeting held on 18 October 1985, the Avionics Panel of AGARD recommended the establishment of an AGARD WG to consider the future potential of Satellite Communications in the NATO context. This proposal was approved by the National Delegates Board (NDB) of AGARD in March 1986 and established WG 13 under the direction of Prof Dr. A.N Ince of Turkey.

OBJECTIVE AND SCOPE

After careful review of the activities of other NATO Agencies addressing SATCOM, the aims of the AGARD WG 13 were agreed to be:

a) To assess emerging technologies in regard to satellite communications and the associated threat environment in the post-2000 era when currently planned systems will be beyond their operational life

b) To develop potential SATCOM system configurations including sub-system design issues

c) From the foregoing work, to identify areas for significant technological thrust and potential roles for satellite communications in the post-2000 era which will assist member states in directing their future R&D programmes as well as NATO in their ongoing development of the NATO C3 Architecture.

The main areas to be addressed are given below *

- Technology assessment
- System configuration concepts (architectures)
- Threat (physical, EW, Nuclear)
- Counter-Counter Measures
- Propagation Issues (including nuclear effects)
- Network and transmission issues
- Satellite payload and launch vehicles
- System management

MEMBERSHIP

The WG will comprise nominated representatives from NATO nations. Participation by NATO Bodies engaged in SATCOM activities is encouraged.

METHOD OF WORKING

The WG will establish at their kick-off meeting a Work Programme and assign individual responsibilities for provision of inputs to the eventual report. Inputs will be coordinated and integrated by the WG Director and the WG will generally meet three times a year to review progress, establish new input requirements and agree the WG Report. The Avionics Panel will be briefed on progress at every Panel Business Meeting.

To achieve the necessary coordination between the work of WG 13 and that of the other NATO Bodies engaged in satellite communications planning, Director WG 13 will keep their Secretaries informed of the progress made.

SCHEDULE

The Programme of Work is expected to last about thirty months. The following are the main milestones:

- Mid 1987	Coordination within NATO and issue of the Terms of Reference and the outline of work areas for the Working Group
- Beginning 1988	First meeting of WG 13
- Mid 1990	Submission of the Final Report to AGARD

SECURITY

The work of the Working Group will normally not exceed NATO SECRET. The need for specific areas of work to be carried out at a higher security level will be considered on a case-by-case basis.

(*) These topics to be addressed were later refined and expanded as shown in Section 4.4

Executive Summary

1. The National Delegates Board of AGARD, upon recommendation by the Avionics Panel of AGARD, approved in March 1986 the establishment of WG-13 to study satellite communications for NATO under the direction of Prof. Dr. Nejat Ince of Turkey.

2. Some 14 scientists/engineers, from research and industrial establishments of Canada, France, The Federal Republic of Germany, Norway, Turkey, the United Kingdom, the United States of America as well as from International Military Staff of NATO and SHAPE Technical Centre, participated in the work of WG-13.

3. The report of which this is the Executive Summary contains fully the results of the studies carried out by the group in the period 1988-1990 on the type of satellite communication systems which NATO can have in the post-2000 era including the critical techniques and technologies that need to be developed for this purpose.

4. In accordance with the Terms of Reference the Group considered a time period beyond NATO IV and other national systems now in the implementation or planning stage, which would cover a time span of 20-30 years, i.e. 2000-2030. It was recognized that the earlier part of this period would be constrained by the existing and planned assets but the later part would be, and should be, more technology-driven. The following assumptions are made which take into account perceived trends and desirable attributes for future SATCOM systems.

   i) The area of interest for NATO will remain as is to-day and will include the polar region(*).

   ii) The use of SATCOM will be more pervasive particularly for small mobile users (aircraft, land-mobile, ships and submerged submarines) to support general purpose and modern C3I structures.

   iii) SATCOM will be integrated with the future ISDN networks now being planned and implemented in the nations and NATO. This may require SATCOM to have improved effective performance with respect to such parameters as delay and echo.

   iv) The need for increased survivability against both physical and jamming threat will continue.

   v) The use of frequencies in the EHF and optical bands for greater capability (e.g. AJ capability and communications with submerged submarines) and smaller terminals are foreseen.

   vi) The future SATCOM systems will be required to be cheaper and more affordable.

   vii) There will be the usual need for interoperability.

5. The Group agreed that the above attributes could be taken as inputs and goals for the system architectures to be developed for a future NATO SATCOM. In fact, these attributes were derived from the deficiencies of the present system which is not flexible enough with respect to growth in capacity and capability, and has a high degree of electronic and physical vulnerability and does not provide communications for the polar region and submerged submarines. The system development in the past has been based on successive discrete steps in capability and spending and each procurement has contained an important cost element of R&D. There has been a minimum of joint national R&D and use of the NATO system which resulted, among other things, in considerable interoperability problems.

   It was agreed that what was required for the coming decades which may be characterized by "uncertainty" and "shrinking military budgets" was a very flexible, modular SATCOM system whose communication capacity and resilience to ECM and physical threat can be modified when operational requirements change, however, without having to undertake excessive R&D and total replacement of the space segment.

6. For the development of system architectures to achieve flexible and highly cost-effective SATCOM options for NATO, a technical survey has been made and information collected on the satellite system concepts being considered nationally (Canada, France, FRG, UK and USA) and internationally (ESA, Intelsat, Eutelsat, INMARSAT) for both civil and military applications as well as on related technological R&D activities and operational aspects regarding threat and environmental factors such as propagation and the usage of frequency spectrum.

7. The status of the following techniques/technologies and concepts which appear feasible and exploitable by future SATCOM systems and which are consequently being investigated nationally and internationally have been described in the report:

   i) multi-beam/phased-array antennas with adaptive spatial nulling and multiple transmit spot beams.

   ii) ECCM techniques.

   iii) flexible and programmable on-board signal processing and switching techniques and devices.

   iv) multi-frequency payloads.

   v) multi-satellite systems to create spatial uncertainty for the enemy.

   vi) use of tethers in space.

   vii) laser and millimetre-wave communications for inter-satellite links,

   viii) blue-green lasercom for submerged submarines.

   ix) application of superconductivity, artificial intelligence, neural networks, robotics and of space-borne computers/software for signal processing and for manual and autonomous control of spatial and terrestrial resources.

   x) power generation in space.

   xi) spacecraft propulsion systems,

   xii) launch vehicles and space transportations.

   xiii) nuclear effects and hardening techniques.

(*) IMS letter dated 15 May 1986 reproduced in Appendix 2C

xiv) physical attack and protective measures against directed energy beams (laser, particle, RF) and ASAT etc.,

xv) sensitive, light, long-life materials, components and devices for sensing, power generation, amplification and control.

8. The speed of progress made in the above areas will be determined mainly by the urgency of the need for, and the amount of resources allocated to, them. These technologies and new production methods coupled with basically software-controlled processing transponders with a capability to continuously adapt to changing requirements are expected to lead to more flexible and reliable, lighter and less-power consuming and altogether more cost-effective satellites than the present ones. Moreover these satellites can be launched by a number of different launch vehicles. Further reduction in cost may be obtained by sharing the satellites (single an/or cluster) between NATO and the Member Countries.

9. It can be stated generally and with confidence that in the time period in question it will be possible to design and build any satillite to meet almost any requirement. Technology exists or will be available for whatever communications performance and level of hardening is required as well as launch vehicles with capability to place the resulting satellite of whatever weight and power into any required orbit. The constraints will be the availability of orbital slots, frequency spectrum and, of course, funds.

10. The cost considerations have therefore been the driving factor for the systems reported here. When assessing different concepts for satellite designs and system architecture what is importanat is not so much their absolute but rather their relative costs. Accordingly a cost model of the satellite system has been established which takes into account:

- frequency band used (SHF, EHF),
- number of transponders,
- spacecraft reliability,
- R&D cost,
- power required,
- weight,
- launch cost,
- recurring cost,
- system availability.

11. Several SATCOM system architectures with the potential of meeting possible future NATO requirements implied in the paragraphs above have been defined using different orbits (geostationary, polar, 12-24-hr inclined at 63°.4 and Low Earth Orbit LEO) and a number of satellites with single and/or dual frequency transponders (SHF and EHF) which can be configured to meet any operational requirement. Table E-1 lists the architectures considered in the report and gives the number of active spacecraft for full continuous coverage of the NATO area including the polar region as well as the total number of spacecraft needed for 7-years and 21-year periods for a certain given spacecraft reliability. Architectures based on the use of LEO and a combination of geostationary and polar-orbit satellites were eliminated from further consideration on cost grounds and the others were subjected to more detailed cost-performance analysis using the cost model mentioned in paragraph 10 above.

12. Table E-2 lists some twenty different promising architectures (cases) for a future NATO SATCOM and gives the associated R&D, recurring and total costs for different spacecraft reliability and continuous service availability for 7 years. The common attributes of these architectures are the following (see Fig. E-1) :

(a) i) The transponders have adaptive receive (with steerable nulls) and multi-beam transmit antennas (1 earth cover, 1 Europe cover, 1 polar spot and 2 steerable spots).

ii) A flexible channelization technique is used on board the satellites at both SHF and EHF. At EHF, this is exploited in an on-board processing concept that prevents the satellite downlink transmitter from being loaded by the jammer and also to prevent unauthorized access to the satellite. For flexible AJ processing and ease of interoperability a full bandwidth (2 GHz at EHF, 500 MHz at SHF) filter band is provided using perhaps different filter technologies to obtain different selectivities required (see Fig. E-1) where the channelization can be controlled by telecommand to avoid interference, to alter the satellite capacity allocated to various geographical areas and to adapt the specific requirements due to restrictions in the tunability of the NATO or national ground segment. At EHF, where the flexible channelization technique is coupled with on-board processing for AJ purposes then switching between high-selectivity filter bank outputs (element filter output) will be performed at a high rate and controlled by an on-board transec equipment which can be programmable (in orbit) to support several simultuneous uplinks.

(b) The ground segment would consist of both SHF and EHF terminals. The SHF terminals would be those existing at the end of the NATO IV era and would be used mainly to support common-user trunks. General transition rom SHF to EHF is foreseen to take place over the period covered in the study to support mainly mobile/transportable users many of which may have demanding AJ and/or LPI requirement.

The EHF ground segment:

i) has preferably non-synchronized frequency-hopped (because of its better performance in disturbed and time-variant propagation conditions and better suitability to small terminals than the direct sequence modulation system) terminals operating in FDMA with flexible data rates and redefinable codes,

ii) consists of the simultaneous accesses (for system comparison purposes) given in Table E-3.

(c) The systems

i) have the virtue of allowing easy transition from existing to future architectures,

ii) Have minimum development, recurring and launch costs,

iii) are upgradable and expandable on a scale to meet operational requirements,

iv) defective and life-expired elements of the system are replaceable without man intervention,

v) spacecraft are capable of being refuelled without man intervention,

vi) have virtually zero down-time at low cost,

vii) make maximum use of orbital slot allocations,

viii) allow spatial distribution of spacecraft to reduce their vulnerability to jamming and physical attack.

It should be noted that the data in Tables E-2 (a) and (b) are for a 7-year period. During the 21-year total period, three stages of complete space segment replacement are expected to occur, which would allow for an update for changes in traffic or other requirements. For dual frequency systems development cost will be incurred at each stage.Single frequency systems will not incur such costs since the same designs of spacecraft would be used throughout the 21-year period; only the mix of EHF and SHF types would change. Provided military components are used in the design of the spacecraft it should be possible to maintain full availabilty over the 21-year interval.

13.An examination of the data shows that:

a) For geostationary operations the cost of interconnection of spacecraft is of the order of 3 % , and is not more than 4.5 % for the inclined orbits (Tundra as the most expensive case).Interconnection provides significant improvement in service availability probability, ranging from 0.14 at the lower inherent spacecraft probabilities to 0.04 at the high end.

b) Increasing the space segment availability by the amount given in (a) above without, however, using Inter-Satellite Links ISL would require launching more satellites and this would increase the system cost by about 25 %.

c) Operation in inclined orbits costs about 50 % more than the geostationary case for the same service availability, but gives full NATO coverage including the polar region.

d) The geostationary case 1 corresponds to the NATO IV satellite as far as coverage and the number of satellites and reliability are concerned. It is interesting to note, however, that the 7-year system cost of Case 1 and that of NATO IV (about 400$M) are almost identical even though Case 1 satellites have considerably more capacity (in SHF and EHF) and significantly greater resistance to jamming (on-board signal processing in EHF and adaptive nulling antennas).

e) The system cost changes significantly with the 7-year service availability probability. How many satelittes would be needed for a 21- year period without having excessive capacity would depend on this as well as on what residual capacity would remain at the end of each seven-year period and how the change in requirements is introduced; abruptly at each 7-year period or progressively during the 21-year period. In the latter case, some reduction in the total number of satellites required and hence in total cost would be expected.

14.The architectures which appear cost-effective and promising are given in Table E-4.

The following comments can be made about these architectures:

a) An adequately wide range oef architectural options are presented from which the architecture best suited to the requirements, as they will be known nearer the date of system implementation, can be selected.

b) Based on the assumptions made regarding possible future NATO requirements, reliabilities of future electronic systems and costs per kilogram of Payloads, Spacecraft Platforms and Launches which have been used consistently for all of the candidate architectures, it can be concluded that:

i) Provided NATO can accept the coverage provided by a geostationary only system of satellites, Architecture A is the lowest cost solution.

ii) If polar coverage obtained by leasing from the USA, costs less than Cost (H-A) or Cost (I-A) then Architectures A or B plus Polar leasing would provide the next lowest cost options. Option B gives improved availability and AJ capability but at 33 % higher cost than the cost of A.

iii) The lowest cost architecture which provides full coverage is Architecture H at a cost increase of 50 % over geostationary only (Case B)

iv) For a further 5 % increase in cost an improvement from 0.95 to 0.98 in operational availability and an enhanced AJ capability can be obtained by using Architecture I. This architecture is probably the most cost-effective option of those considered. to all of the assumed future NATO requirements.

15.It is likely that cost will be the driving factor in determining the hoice of a future SATCOM architecture and it is therefore appropriate to consider the three dominant cost factors (R&D. replacement and launch costs) and indicate what steps could be taken to bring about cost reduction in each case.

a) Economy in R&D costs could be obtained through NATO/National collaboration and by adopting a modular approach to system diversification and evolution. An effective way of achieving the latter would be to develop at the outset separate SHF and EHF spacecraft and use them, in GEO and TUNDRA orbits alike, in a mix determined by the changing requirements.

b) The use of smaller spacecraft, even though more of them may be needed, could lead to lower system costs because of the economies of scale. Such economies of scale would be further enhanced if the same spacecraft types are used at all stages of system evolution over a period of, say, twenty years. They will also be enhanced if the same spacecraft types are bought for national as well as NATO use.

c) Launch costs can be minimized by reducing spacecraft mass, in particular through the exploitation of new technology. It is also important to maximize compatibility with the largest possible range of launch vehicles.

d) Interconnection of spacecraft increases system reliability and therefore tends to reduce the total number of spacecraft that need to be launched.

e) Finally, long-term planning is the key to achieving reductions in both R&D and recurring costs.

16.The NATO SATCOM systems so far acquired have been based on national developments adapted to NATO requirements and the continuity of service (not necessarily full service) has been obtained by sharing or borrowing capacity from national systems. The national systems, in turn, have relied for continuity of service on the availability of capacity on the NATO system. Each procurement has contained an important element of R&D costs and since successive systems have been developed almost independently of each other, R&D costs have been, like the replacement cost, also recurring.

There has been a minimum of joint national R&D and use of the system and each procurement has been preceded by lengthy negociations on production sharing which has not, in general, satisfied, at least, some of the member countries. As a result of having independent NATO and national systems there has been considerable interoperability problems.

It is believed that this trend, based on successive jumps in spending and capabilty with a minimum degree of general national participation should be and can be changed to meet the needs of the coming decades which may be characterized by uncertainly and shrinking military budgets requiring affordable and flexible systems.

The member countries have adequate experience within NATO and Europe and know that under these circumstances it is necessary to resort to joint R&D, procurement and use of the system while ensuring effectiveness and competitiveness for keeping the costs down.

17.What needs to be done jointly are:

a) To define NATO and national requirements for satellite communications.

b) To develope and agree on a system architecture.

c) To delineate those technological areas which are critical and require R&D.

d) To encourage and support companies and R&D establishments to form research partnership for development and production to be carried out in a competitive manner.

It is believed that the architectures evaluated and recommended *in this report* form a good foundation for (b) above and ensure also that the satellite designs outlined that are flexible need not change basically over a period of some twenty years or longer thus keeping the R&D and recurring costs to a minimum.

The report outlines also certain critical technologies for (c) above which need R&D. Some of these R&D topics are common to military and civilian satcoms, some others which are specific to military, are likely to be common to both national and NATO systems and yet another category of topics will be NATO-specific. It would therefore be necesarry to make a more detailed assessment of the R&D topics and determine where R&D is a prerequisite and either can be relied upon present/future civilian developments or carried out jointly by member nations.

18.The following areas appear as first candidates for a NATO R&D effort because they would provide solutions to problems which are NATO-specific and can be made available within the time-frame considered for NATO SATCOM systems:

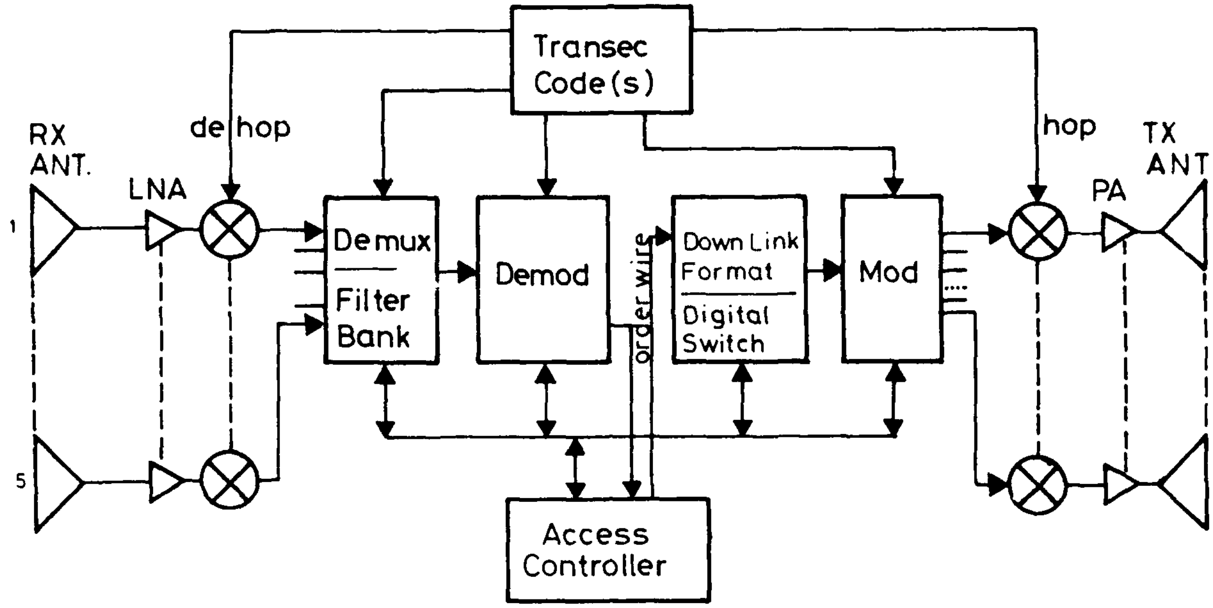
a) On-board flexible anti-jam signal processing which can be controlled by software to meet AJ requirements generally and to adapt to national modems and new modems introduced during the lifetime of the satellite(s).

b) Adaptive nulling AJ receive antennas tailored to NATO needs and to keep costs down.

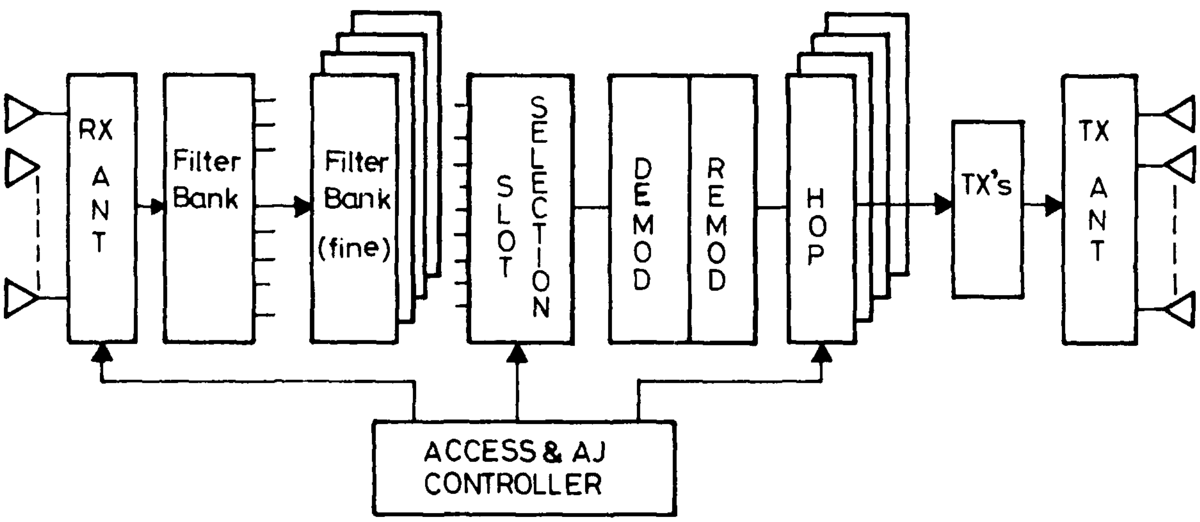
c) Autonomous control of the spacecraft and O&M generally using techniques of artificial intelligence, neural networks and robotics.

By sporing R&D activities, limited to payload technology, in the member nations even on modest scale, NATO can expect to get a better insight into and to make an impact on current technology developments carried out for civilian and military purposes.

19.It is believed that the collaborative approach outlined above for SATCOM acquisition in NATO give tasks to all the existing bodies in NATO such as NACISA, STC, DRG, IEPG, AGARD, etc., and would probably not necessitate creating new structures. Cost and benefit anlysis carried out in the report show that the architectures recommended can be implemented in the manner suggested above and could lead to systems which are considerably cheaper and much more effective than those we have had so far.



a) Non-transparent satellite with about 50 MHz information bandwidth hopped across 2000 MHz. All terminals hop in synchronism. There will be as many dehoppers in the satellite as there are different nets with different codes.



b) Transparent satellite (full bandwidth) allowing operation with non-synchronised frequency-hopped terminals.

Fig. E-1 Two promising ∴J satellite designs with on-board processing and routing

Table E-1
Number of satellites required for different SATCOM architectures

Features Architecture	No of Active S/C for full continuous Coverage	Comment	No of S/C for 7 years	Total No of S/C for 21 years
(a) Proliferated LEO	240		>240	>500
(c) Inclined Elliptical Orbit (c1) LOOPUS (c2) MOLNIYA (c3) TUNDRA	 9 3 x 12 - hrs 2 x 24 - hrs 2 x 24 - hrs		 27 9 6 6	 81 27 18 18
(d) GEO	1	Baseline	3	9
(e) Systems of Satellites in more than one Orbit (e1) GEO + Polar	 1 GEO + 6 Polar	 S/C in GEO are dual frequency single in Polar orbit	 3 GEO + 18 Polar	 9 GEO + 54 Polar
(e2) GEO+24-hr MOLNIYA (a) Dual freq S/C in GEO+single in Inclined Orbit (b) Single freq in both orbits	 1 GEO + 2 inclined 2 GEO + 2 inclined	 GEO S/C have EHF and SHF Inclined EHF or SHF GEO S/C are air of EHF and SHF Inclined EHF or SHF	 3 GEO + 6 inclined 6 GEO + 6 inclined	 9 GEO + 18 inclined 18 GEO + 18 inclined
(f) CLOUDSAT with 10dB ECCM advantage relative to (e2)b	Receive S/C 20 GEO. 20 Inc. TX S/C:- 2 GEO+2 Inc	Numbers will depend on jamming threat at the time	60 GEO 60 Inclined 6 GEO 6 inclined	180 GEO 180 inclined 18 GEO 18 inclined
(g) MEWS	The basic concept applies to all case (c) thro' (f)			

Note : - This table is valid with the assumptions given in Section 10.6.8.

Table E-2 (a)
System evaluation data for geostationary operations (7 years)

Case	Spacecraft		Payload			Availability		Total No. Operating Spacecraft	System Costs			
	No / Orbit	Frequency	EHF	SHF	Connect	S/C	Space Segment		R & D	Recur	Launch	Total
1	2	Dual	Y	Y	No	0.61	0.85	2	242	82	76	400
2	2	Dual	Y	Y	Yes	0.61	0.93	2	262	87	79	428
3	3	Dual	Y	Y	No	0.61	0.94	3	242	163	115	520
4	3	Dual	Y	Y	Yes	0.61	0.98	3	262	173	118	553
5	4	Single	Y	Y	No	0.72 0.76	0.87	4	290	96	92	478
6	4	Single	Y	Y	Yes	0.72 0.76	0.93	4	321	106	96	523
7	6	Single	Y	Y	No	0.72 0.76	0.96	6	290	192	138	620
8	6	Single	Y	Y	Yes	0.72 0.76	0.99	6	321	212	144	677

Table E-2 (b)
Evaluation data for 24-hour Tundra orbits (7 years)

Case	Spacecraft		Payload			Availability		Total No. Operating Spacecraft	System Costs			
	No / Orbit	Frequency	EHF	SHF	Connect	S/C	Space Segment		R & D	Recur	Launch	Total
1	2	Dual	Y	Y	No	0.61	0.72	4	223	221	126	570
2	2	Dual	Y	Y	Yes	0.61	0.86	4	238	236	130	604
3	3	Dual	Y	Y	No	0.61	0.88	6	223	368	189	780
4	3	Dual	Y	Y	Yes	0.61	0.96	6	238	393	195	826
5	4	Single	Y	Y	No	0.72 0.76	0.76	8	258	255	155	668
6	4	Single	Y	Y	Yes	0.72 0.76	0.87	8	288	285	163	736
7	6	Single	Y	Y	No	0.72 0.76	0.93	12	258	425	233	916
8	6	Single	Y	Y	Yes	0.72 0.76	0.97	12	288	475	245	1008
9	1	Single	Y	-	No	0.72	0.52	2	103	34	33	170
10	2	Single	Y	-	No	0.72	0.85	4	103	102	66	271
11	2	Single	Y	-	Yes	0.72	0.93	4	118	117	70	305
12	3	Single	Y	-	Yes	0.72	0.99	6	118	196	104	418

Table E-3
The EHF Ground segment assumed for system comparison

Type of Terminal	Antenna Dia. (m)	Tx Powers (kW)	Transmission Rate (Baud)	Number of simul-taneous accesses
Ship- Borne	1.0	0.1	2400-9600	20
Aircraft	0.5	0.1	2400	10
Submarine	0.25	0.1	100-2400	1
Land transportable	5.0	1.0	4 x 64000 (*)	15
	2.0	0.5	4 x 2400	30
Man-pack	0.5	0.01	75	30

(*) 16 kb/s codecs (adaptive sub-band coding) exist today with quality which equals that of the 64 kb/s PCM. It is expected that in the timeframe considered in this report there will be 8 kb/s or even lower-rate codecs available for use in SATCOM with qualities comparable to that of 64 kb/s PCM voice (see Section 6.3).

Table E-4
Cost-reliability comparison of candidate architectures for a system lifetime of 21 years

Architecture	Cases No		Total No of operating spacecraft	No of Payload		Orbit / Freq.		Service Availability	Cost (21-year) ($M)			
	T	GEO		EHF	SHF	T	GEO		R & D	Recur.	Launch	Total
A		1	2	2	2	-	D	0.85	3 x 242 = 726	3 x 82 = 246	3 x 76 = 228	1200
B	-	4	3	3	3	-	D	0.98	3 x 262 = 786	3 x 173 = 519	3 x 118 = 354	1659
C	2	-	4	4	4	D	-	0.86	3 x 238 = 714	3 x 236 = 708	3 x 130 = 390	1812
D	4		6	6	6	D	-	0.96	3 x 238 = 714	3 x 393 = 1179	3 x 195 = 585	2478
E	6	-	8	4	4	E,S	-	0.87	288	11 x 95.01 = 1045	12 x 40.8 = 489.4	1822
F	8	-	12	6	6	E,S		0.97	288	1615	734	2637
G	12	4	9	9	3	E	D	0.97	118 + 786 = 904	665 + 519 = 1184	31[illegible] 667	2755
H	12	7	12	9	3	E	E,S	0.95	290.5	1470.2	727.1	2488
I	12	8	12	9	3	E	E,S	0.98	320.86	1550.9	745.3	2617
J	8	8	18	9	9	E,S	E,S	0.96	320.86	2558.2	1166.1	4045

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

At their Business Meeting held on 18 October 1985, the Avionics Panel of AGARD recommended the establishment of an AGARD WG to consider the future potential of Satellite Communications in the NATO context. This proposal was approved by the National Delegates Board (NDB) of AGARD in March 1986 and established WG 13 under the direction of Prof. Dr. A.N. INCE.

Recognizing that there were a number of bodies in NATO who dealt with different aspects of satellite communications, it was agreed by all concerned that there was a need for care to avoid unnecessary duplication of work in this area within NATO.

The first coordination meeting chaired by Prof. Ince, took place on 23-24 June 1987 at the NATO Headquarters. The representatives appointed by seven NATO countries for the WG 13 as well as those from NATO IMS/CIS Division, NATO CIS Agency (NACISA) and SHAPE Technical Centre attended the meeting. Also in attendance was Dr. HUNT, Chairman of the Avionics Panel who briefed the meeting on AGARD and AVP generally and on the Panel's space interest in particular. He stated that following the NDB decision, Prof. Ince and the Panel requested inputs and comments from NATO bodies engaged in SATCOM work. He stated that answers were received from the NATO IMS, STC, and the Director C3 of NATO. It had also been discussed in the Panel and by the appointed members for the WG 13. The purpose of the meeting was to find a way ahead for the Group to function, to that end the original TOR would be reworked to take into account the written and verbal comments received from the NATO Bodies mentioned above.

As a prelude to discussions on the subject the representatives of the NATO Bodies briefly described the work being carried by them which relate basically to the problems of the existing SATCOM system and its future architecture planning. NACISA is charged with managing the development of NATO communications including SATCOM. The Satellite Communication Working Group (SCWG) provides policy level advice to the NATO communications and Information Systems Committee (NACISC) on issues of system concept, scope, schedule and performance for future NATO SATCOM. The Tri-Service Group on Communication and Electronic Equipment (TSGSEE) under the Commitee of National Armament Directors (CNAD) is charged with the development of interoperability standards for satellite ground terminals. SHAPE Technical Centre (STC) gives technical support and contributes to the work of all these bodies.

The meeting then discussed all the issues involved, and took note of the Director's proposal that the work to be undertaken by WG 13 should be technologically driven (and not requirements driven) and should develop potential SATCOM system configurations including sub-system design issues and identify R&D areas and potential roles for SATCOM in the post-2000 era which could be helpful for the member states as well as for NATO for their future R&D programmes and SATCOM planning work.

The meeting commenting that:

i) the SATCOM studies to be undertaken by WG 13 should be technology driven and should include considerations of potential system configurations and roles for SATCOM in the post-2000 era

ii) the time period considered should be such that the studies would not be undully constrained by the past and yet it would be possible to make sensible and reliable forecasts for the feasibility and viability of the technologies needed by the systems to be defined,

iii) it should be assumed that technologies available to NATO would also be available to the enemy, agreed on the "Terms of Reference" as reproduced at the beginning of this report. This TOR as well as the work assignments made to the members at the meeting were later approved formally by all the parties concerned. It was also agreed that WG 13 would keep informed all the NATO Bodies engaged in SATCOM planning work on the progress made by the Group.

1.2 TERMS OF REFERENCE

The terms of reference for the Group were reviewed at the first technical meeting held on 22-25 February 1988 with a view to better defining the time-period to be covered by and the scope of, the studies.

After a thourough discussion it was agreed that WG-13 should consider a period beyond NATO IV and other national systems now in the implementation or planning stage. The period for studies should therefore correspond to NATO V and VI programs and cover a time span of 20-30 years, i.e. 2000-2030. It was recognised that the earlier part of this period would be constrained by the existing and now planned assets but the later part would be, and should be, more technology driven. The group would not deal with presently conceived operational requirements and system architectures but it was stressed that, even for a technology-driven study, it was essential to have in mind an objective and functions to be fulfilled by future SATCOM systems.

In passing it was noted that, unlike some NATO bodies, WG-13 was an advisory body and had no line responsibility and consequently it had the freedom of examining any aspect of SATCOM which appeared potentially useful to NATO. It was further agreed that duplication was not necessarilly bad since two independent groups would often come up with different conclusions/solutions even when they study the same problem.

In connection with future requirements it was pointed out that The Avionics Panel of AGARD ran in Oct. 1985 a workshop on "The Potential Impact of Developments in Electronic Technology on the Future Conduct of Air Warfare" which resulted in AGARD Advisory Report No.232 containing information on communications requirements which cannot be met by present day systems and on the state of the art and future trends in technology which may make powerful satellite communications systems possible. It was agreed that the deficiences thus established could form the basis of the functions to be performed by the future systems.

1.3 METHODOLOGY

The method of working to be used was agreed as follows:

i) The present members and those who could not attend would form the core staff of WG-13. Each member would be responsible for the work assigned to him which he

either would carry out himself or get it done by others in his country/establishment etc. observing always the applicable security rules.

ii) The work assigned in a meeting would be completed by each member and sent to the director about a month ahead of the following meeting. The director would then edit/collate the information received adding his own contribution and send the resulting report back to the members. The report would then be discussed and agreed by the members at their next meeting when they identify and assign to themselves new work items. The director would keep the Panel informed of the progress made and receive from them guidance about the future conduct of studies.

iii) The members would do their best to keep themselves and the group informed of the national and internationl developments in the satcom area.

iv) The members would try to keep the security classification of the documents/reports they produce down to "NATO unclassified of restricted", if possible at all and where necessary omitting or deleting some sensitive information which can be filled in at the meeting etc.

v) There would be three meetings of the WG a year held at venues which may be:

a)at NATO HQs and other NATO Agencies where easy access to NATO documents and security arrangements may be obtained.

b)at National establishments where technical visits may be made and,

c)coincident with the venue of AVP symposia providing convenience and economy in cost and time for the WG chairman and other panel members who have to attend the symposia for other reasons.

vi) In order for WG-13 to produce work relevant to NATO and the nations and to avoid unnecessary duplications, it was agreed, it was necessary to have information not only on NATO programs but, of course, also on national and international programs including those of Canada, France, FRG, the United Kingdom as well as ESA projects and the NASA, US MILSTAR, SDI programs as having great potential for influencing the future course and development of space systems and satellite communications. It was recognized that there would necessarily be restrictions on the release of information on some aspects of the national plans and programs but the members would do their best to collect and submit to WG-13 information that is releasable to NATO.

vii) The following reports, documents, symposia and meetings would be used to extract information relevant to the work in hand:

a) AGARD Advisory Report No: 232, 1985
"The Potential Impact of Developments in Electronic Technology on the Future Conduct of Air Warfare".

b) Special Issue of Proc. IEEE, Nov. 1984 on
"Satellite Communication Networks"

c) IEEE Journal on Selected Areas in Communications
"Satellite Communications Toward the Year 2000"
May 1987, Vol. SAC-5, No:4.

d) Special Issue of Proc. IEEE, Jan 1987.
"Packet Radio Networks".

e) A report dated Nov. 1985 by the American Institute of Aeronautics and Astronautics submitted to the National Commission on Space dealing with the technology base that would be critical to possible civil space endeavors in the 2000-2030 time period.

f) A paper based on a study (funded by NASA) by the Aerospace Corporation entitled "Advanced Space System Concepts and their Orbital Support Needs" outlines "Innovative Applications of Very Large Satellites Communicating with a Myraid of Tiny Inexpensive User Terminals".

g) The NATO SATCOM Interoperability WG documents:
NACISA/NSIWG(86) 1 and
NACISA/NSIWG(86) 2

h) SATCOM related symposia:
- "Military Applications of Space"
by AAAF/France, Oct. 1988
- "Military Microwaves"
London, 5-7 July 1988
- IEEE MILCOM 88, San Diego, Oct. 1988

i) A study conducted by Sub- Group 21 of NIAG on tactical communications, called TACOM 2000.

1.4 CONTENT OF THE REPORT

To set the scene for developing possible future SATCOM architectures, aimed at correcting existing deficiences and providing capabilities to meet future perceived requirements, Chapter 2 reviews the present NATO communications network and the evolving "NATO C3 Architecture" and discusses the implications on and the future role of SATCOM in NATO including the need for communications in the polar region and for submerged submarines as well as for protection against nuclear effects.

Chapter 3 attempts to identify and analyse both the similarities and the differences between NATO and other civilian/military SATCOM systems in terms of the type of services provided and the environmental factors. The objective is to determine the areas in which NATO can draw direct benefit from the R&D work done for the development of other systems so that NATO specific areas may be delineated for further consideration and action by NATO.

In Chapter 4 all the issues and trends in satellite communications are listed and the responses by the Group are given. Some of the views expressed were used in determining the areas on which the present studies were to focus and in drawing up a detailed program of work for the Group taking into account also the Terms of Reference.

One of the most important drivers of SATCOM system design and cost is the threat the system is expected to face, in the time-period of its operation, and survive wholly or partially. The perceived thereat to NATO communications in the post-2000 era is treated in Chapter 5 in terms of jamming of the up and down links, interception, piracy, as well as physical and nuclear attacks.

The implications of the threat postulated in Chapter 5 are analysed in Chapter 6 where the concept of survivability as applied to ECM/ECCM and the different types of spread spectrum and coding techniques are discussed followed by a review of hardenning measures which may be taken to protect the satellite against effects of physical and nuclear attacks.

The environment in which a SATCOM system has to work can affect the signal attenuation and fading which are highly frequency dependent and this subject is treated in Chapter 7 which considers all propagational factors (ionisation, precipitation, particulate matter, hydrometeors and reflections) and discusses their effects on EM propagation in terms of absorption, scattering, multipath, scintillation, and depolarization. The important subject of noise and mutual interference between systems, particularly those using the geostationary earth orbit is also reviewed. The implications of propagational factors, interference and noise are evaluated and measures which may be taken to counter the undesirable effects are treated in this chapter with particular emphasis on the EHF band which is of great interest to NATO.

Chapter 8 is where a review is made of the important field of technology which is considered relevant and assential for the development of the type of future SATCOM systems postulated and defined in the report. The range of techniques and technologies covered include devices for analog and digital signal processing, adaptive antennas, laser intersatellite links, lasercom for submarines as well as launch vehicles, power generation in space, new materials and the application of artificial intelligence, neural networks and robotics for autonomous system operation and maintenance. This chapter concludes with recommended principles which may used to select R&D projects for support by NATO.

The NATO and national systems and developments, again with relevance to the work of the Group, are described in Chapter 9 which is based on information provided by the members of the Group and also obtained from presentations given and laboratories visited in connection with the meetings of the Group, held in various member countries.

Generic technical charactestics of a future NATO SATCOM system based on Chapters 2,4,9 and the technical possibilities offered by techniques/technologies discussed in Chapter 8, are postulated in Chapter 10 which then proceeds to develop, define and evaluate several interesting system architectures from which three basic candidate architectures for NATO are derived. The number of satellites required with full coverage of the areas of interest to NATO and service continuity for a system of 21- year lifetime for different SATCOM architectures is also given in this Chapter.

The candidate architectures outlined in the previous chapter are compared in cost and performance in Chapter 11, based on defined ground segment and satellite features, assumed to be the same for all cases, and using a cost model which is also described. The model takes into account payload, frequency band, orbit type, mass, power, reliability, and R&D costs which are all detailed for some twenty different cases.

Chapter 12 is where a new method of acquisition for future NATO SATCOM systems is suggested. This is based on a collaborative approach between NATO and the member countries undertaking certain actions jointly in defining requirements, developing the system architecture, delineating and supporting critical R&D areas and finally using the system jointly.

Finally, in Chapter 13 conclusions and recommendations are given pointing out the advantages for NATO of the candidate architectures described in Chapter 11 and the areas of new technology which NATO should support if it wishes to implement a future NATO SATCOM system based on the recommendations of this report.

CHAPTER 2

CURRENT NATO COMMUNICATIONS PLANNING (SATCOM AND THE NATO C3 ARCHITECTURE)

2.1 PLANNED EVOLUTION OF THE NICS

Over the next two decades, communications within NATO are to undergo a fundamental change. At present the NICS is implemented using NATO-owned bearer and switching facilities and rented PTT circuits, but is planned to phase these out in favour of a scheme whereby NATO will rely principally on National military communication networks to meet its needs. This will be facilitated by the widespread introduction of digital switching technology based on CCITT standards. In the near term, NATO will provide gateways and cross-border links to form a robust system of inter-connected NATO and national networks. NATO subscribers will be progressively transferred to national digital networks and NATO networks such as IVSN and TARE will eventually cease to exist as separate entities. Provision of a common user interface based on the Integrated Services Digital Network (ISDN) concept will allow all existing and foreseen NATO subscribers' voice and data requirements to be met. To the greatest extent possible, NATO users will be served by this common-user system. However, it is recognised that independent "emergency" communications systems will probably still be required to provide highly survivable circuits to vital users in tension and war. Such systems must be capable of interfacing with the common-user network. The totality of NATO communications systems, common-user and emergency, will form the "NATO C3 Architecture".

Two developments have prompted this change in philosophy. Firstly, the ACE-High system of microwave and tropo-scatter links which has supported many of the NICS trunks since the 1960s has reached the end of its useful life and has come up against serious frequency-coordination problems. It is being replaced by the NATO Terrestrial Transmission System (NTTS) consisting of all-digital cross-border links. Secondly, the widespread introduction of digital switching and high-capacity fibre-optic bearers promises to give national networks adequate connectivity and capacity to meet NATO's communication needs in addition to national needs.

It is intended to indroduce ISDN standards at an early stage in the evolutionary process. According to current thinking [2.1], the internetwork gateways will form switching nodes within the core of a "NATO ISDN", the concept of which is illustrated in Fig. 1. The inter-gateway links will be supported by a mix of NTTS, NATO SATCOM, and digital circuits established through nationally-owned sub-networks. Initially, the majority of NATO users will access the NATO ISDN via "Access Networks", each of which could be a national military network or an exiting NATO network such as IVSN. "Access Interfaces" will be provided to implement the protocol and signalling conversion needed between the Access Network and the NATO ISDN. A-to-D and D-to-A conversion will also be perpormed here if necessary, as will data rate conversion. In the long term, most national military networks are expected to adopt ISDN standards, making the full range of ISDN services available to NATO users. This will also facilitate network interconnection, though it is likely that the NATO ISDN will be preserved so as to police and optimally route inter-network calls.

The ISDN concept has evolved to meet commercial needs and therefore does not provide all the facilities required of a military communications system. In particular, it does not provide encryption, Traffic Flow Security, precedence or pre-emption. It is envisaged that as far as possible these additional functions will be provided at the boundary of the NATO ISDN, for example by "Virtual Terminals" implemented within the Access Interfaces or inter-network gateways.

Principles of ISDN are summarized in Appendix 2.A where some possible future developments are also discussed.

2.2 IMPLICATIONS OF THE C3 ARCHITECTURE FOR NATO SATCOM

The implications of the transition to an ISDN-based architecture for NATO SATCOM are still being determined. However it seems possible that there will be an increase in the demand for SATCOM in the early stages, since SATCOM will be required to provide some of the bearers in the NATO ISDN while still supporting the IVSN and TARE as "Access Networks". This will be offset by the progressive removal of selected IVSN and TARE SATCOM links as the subscribers they serve are transferred to national networks.

ISDN offers the user both digital voice and data facilities. These are available simultaneously if required. Speech is encoded using 64 kb/s PCM. The standard "2B + D" interface provided to subscribers comprises two 64 kb/s channels and one 16 kb/s channel for signalling and packet data. Within switched network capacity is normally provided in multiples of 64 kb/s. End-to-end transparency is a basic principle of ISDN.

It seems likely, then, that the adaption of ISDN standards will imply a phasing-out of 32 kb/s CVSD SATCOM circuits as IVSN subscribers transfer to national networks. At the same time there will be an increasing demand for 64 kb/s SATCOM bearers for the NATO ISDN. The number and connectivity of such bearers is as yet unknown, but is unlikely that the total number of point-to-point SATCOM circuits provided via the NATO inter-static network can exceed that currently supported even if all SGTs are broad-banded, since the increase in bandwidth of the NATO IV satellite compared with that of NATO III will be taken up by the transition from 32 kb/s to 64 kb/s links. In the post-NATO IV era, it may be possible to accomodate additional common-user circuits at SHF through frequency re-use, but this will have to be shown to be cost-effective vis-a-vis use of alternative communications media.

When jamming is present, NATO SATCOM will fall back to an ECMM mode of operation using spread-spectrum modems to support the more essential users. Data rates must be minimised in this mode so that each terminal can provide an adequate signal-to-jammer ratio for as many users as possible. Thus, for example, 2.4 kb/s vocoded speech would be used in place of 64 kb/s PCM. It may be possible for links within an ISDN to revert to such an ECCM mode without violating basic ISDN principles on the basis that "information transparency" rather than "data transparency" is the requirement for voice circuits. However, this is arguable and a subject for further study. The alternative approach is to treat ECCM SATCOM as a separate non-ISDN system which can be accessed from the NATO ISDN through suitable interfaces. ECCM data circuits may more easily be accomodated within the ISDN concept since reversion to a low rate is possible without losing bit integrity, provided adequate buffering and flow control is provided within the network.

A number of technical problems will arise when SATCOM bearers are used in a NATO ISDN (see Appendix 2B). Several of these are due to the propagation delay. This means that the signalling protocol (CCITT No. 7) used on terrestrial links will not work on links via geosynchronous satallites. A suitable protocol for such links has been specified, but protocol conversion between the terrestrial and satellite standards will be needed at each SGT.

Although the CCITT recommendations for ISDN allow for single-hop satellite delays, multiple hops are barred and the system control will have to prevent these. Another difficult area is network timing: SATCOM modems will have to synchronise bit timing with that of the ISDN, and provide buffering to absorb variations in propagation delay while maintaining bit integrity.

In the long term, interconnected national military networks sharing common standards will provide a highly connected common-user system for NATO, at least in peacetime and tension. Use of fibre-optic trunks will ensure adequate capacity, excellent circuit quality and immunity from interception. As the common-user system evolves towards this goal, it may become increasingly difficult to justify continued use of a "permanent" SATCOM overlay: The high cost of maintaining and manning the terminals of the NATO inter-static SATCOM network may well make it no longer cost effective to continue using this network as a complementary transmission medium to the terrestrial bearer system. Where SATCOM will have a part to play is in war when the common-user network has become fragmented. Transportable SATCOM terminals can be rapidly deployed to establish survivable ECCM links, albeit of low capacity, between ISDN "islands" in an effort to restore full connectivity. And of course, SATCOM will continue to provide access links into the commen-user network from mobile users such as ships.

2.3 PLANNED EVOLUTION OF THE NATO SATCOM SYSTEM

2.3.1 Introduction

The evolution of NATO SATCOM up to the end of the century (but not beyond) is already planned to a large extent. The Tri-MNC C2 Plan (Edition 4) indicates the cost and timing of planned enhancements to the system and of the introduction of new elements. The NATO SATCOM Enhancement Working Party (NASEWP) has made additional recommendations [2.2] which, if endorsed, would be rationalised with the C2 Plan. However, all such planning is likely to come in for critical review in the light of the recent political developments in Europe.

In the next decade NATO SATCOM will continue to support common-user bearers via the large static SGTs, but there will be much more emphasis on the provision of highly-survivable SATCOM links to support vital users in tension and war. This will be made possible by exploiting the "military" features of the new NATO IV space segment. The main system enhancements planned up to 2000 are summarised below.

2.3.2 NATO IV Space Segment

The launch of the first NATO IV satellite is planned for December 1990. Two satellites are being procured, in order to provide adequate availability for a ten-year system life. Like NATO III, the satellites will be placed in quasi-geostationary orbit with the operational satellite at 18°W longitude. The orbit will be inclined at 3.5° initially to avoid the need for North-South station keeping. Nato IV is three-axis stabilised satellite based on the UK Skynet 4 design. It is built by British Aerospace with Marconi Space Systems as cocontractor with responsibility for the communications payload.

Like NATO III, the main feature of the payload will be wideband, limiting SHF transponders, but there will be four of these in NATO IV with a total bandwidkh of 340 MHz as opposed to three transponders covering 150 MHs for NATO III. Each transponder is associated with a different coverage area, viz. NATO Region, Global, European and Central European spot. The total EIRP of the SHF transponders is about 40 dBW compared with 35 dBW for NATO III. Uplinks are normally received via a global coverage antenna, but one of the transponders can also receive via the Central Eurapean spot beam antenna if required.

The SHF repeater incorporates a number of special features to improve system performance in the presence of jamming.

The NATO IV payload also includes a UHF repeater, unlike earlier NATO satellites which have been entirely SHF. The repeater provides two narrowband (25 kHz) channels with global coverage, but has no ECCM features. Although NATO has no formal requirement for UHF SATCOM, indications are that demand from nations for use of the channels-principally for maritime applications-will be great, and that some form of Demand-Assigned Multiple Access (DAMA) will be necessary in order to make best use of this limited resource.

2.3.3 Enchanments to the Inter-Static Network

No major extension of the inter-static network is forseen in the next decade, but a modest growth in the volume of common-user traffic is expected. Certain modifications to the SGTs will be necessary if the additional capacity of the NATO IV satellite is to be exploited to cater for this increased traffic. At present the SGTs use klystron HPAs with a (tunable) bandwidth of about 200 MHz. Thus each SGT can access only about 50 % of the NATO IV transponder bandwidth at any one time. Access plans have been developed [2.3] which, with suitable tuning of the indivdual HPAs, allow the existing inter-static traffic load to be supported on NATO IV, but these have no scope for expansion. Only by increasing the instantaneous bandwidth of the individual SGTs - for example by fitting new TWT HPAs - can this limitation be overcome. It seems likely that most, if not all, of the large static SGTs will have been broad banded in this way by the end of the century but a detailed timescale has yet to be agreed.

2.3.4 Support of Medium Static Terminals

The NATO IV system will support links to and from a number of medium (7-metre) static SGTs in fulfillment of new NATO and National requirements. Two such SGTs are being provided in support of the Improved UK Air Defence Ground Environment (IUKADGE). They will be located in the Shetland and Faroe Islands and each will transmit medium-rate radar data via one of the large static SGTs. Four medium SGTs will be procured for the so-called "Island Command" role. These will be sited in Greenland, the Azores, Madiera and Bermuda. Effectively they will extend common-user network facilities to these sites. A further medium SGT may be installed in West Crete.

2.3.5 Mobile Terminals

Perhaps the most important innovation to the Ground Segment in the next decade will be "Highly Transportable Terminals" intended to provide survivable SATCOM circuits for a number of purposes in tension and war. These SHF terminals will rely on mobility and Low Probability of Intercept (LPI) characteristics to avoid detection and will be provided with ECCM modems to overcome uplink jamming. They will have electromagnetic pulse (EMP) protection. The specifications of these terminals are currently being drawn up. Two sizes have been suggested: A 6-metre version and a 2.4-metre version.

Individual HTTs may be required to support more than one emergency network, and therefore some form of demand assigment which takes account of user precedence will be needed. Principal circuit requirements will be for 2.4 kb/s secure speech, medium/low speed data and 75 b/s telegraph.

Proposed roles for the HTTs include Crisis Management and Political Consultation (CMPC), support of mobile Alternate War Headquarters (AWHQ) and the provision of emergency communications for Theatre Nuclear Forces (TNF).

Because of funding limitations, introduction of the HTTs is likely to take place in piecemeal fashion over a prolonged period. Urgent requirements such as the SHAPE AWHQ will have priority and will be met in the early 1990s. However, it is unlikely that the

full complement of HTTs as currently envisaged will be operational much before the end of the century.

As distinct from the HTTs, two 1.8 -metre mobile terminals have been procured for the ACE Mobile Force (Land). These terminals are built to the UK's VSC 501 specification and will be fitted with the VSC 330 ECCM modem. This modem will also be installed at selected large static SGTs which will act as anchor stations.

2.3.6 Shipborne Terminals

NATO SATCOM will continue to provide communications for national ships when under NATO command throughout the NATO IV era and also, under MOUs, to nations such as Germany and the Netherlands which do not have military communications satellites of their own. The main requirement will continue to be for low-rate data and voice ship-shore-ship communications via designated SGTs. At SHF, CDMA modems will gradually be replaced by more ECM-resistant systems such as OM-55. Maritime links will normally be restricted to the 50 MHz- wide "exclusive band" between 7925 and 8025 MHz.

Unlike previous NATO satellites, NATO IV will carry a UHF transponder and there is considerable interest in the use of this for naval applications. It is likely that DAMA will have to be introduced in order for all potential users to have access to the two narrowband channels available. UHF SATCOM is already extensively used by the US Navy and has the advantage of simplicity in the shipborne equipment. However, its vulnerability to both jamming and unintentional interference is well known.

2.3.7 NATO Air Base SATCOM System (NABS)

Under a NATO/US MOU, the US are providing a number of medium (6-metre) SGTs at US Air Bases which will be supported by the NATO SATCOM space segment. A total of about 80 such terminals will be deployed in two phases: In the first phase, beginning in 1991, 54 terminals will be installed. These will be equipped with Harris Type MD-1131 ECCM modems. The second phase SGTs, whose installation is not expected to begin before 1995, will be equipped with the Universal Modem (see below). Phase 1 terminals may be retro-fitted with this modem eventually.

Some of the sites being equipped with NABS terminals are among those requiring SATCOM facilities for TNF emergency communications. Possiblities therefore exist for rationalisation and these are currently being investigated.

2.3.8 Support of NADS/IADS

Under another US/NATO MOU, NATO will provide SATCOM capacity to support US Air Defence communications requirements around Iceland. The exiting large static SGTs will anchor links from US shipborne and mobile terminals and support national circuit requirments between Iceland, the USA and Europe. The effect is a relatively small increase in the overall traffic load of the NATO SATCOM system.

2.3.9 The "Universal Modem" (UM)

As the result of a US-UK initiative, early development work has been carried out on a new ECCM SATCOM modem intended for use in the mid-1995 and beyond. The modem uses frequency-hopping and has features to mitigate the effect of exo-atmospheric nuclear bursts as well as to combat "intelligent" jamming strategies. At the time of writing, "brassboards" have been built by two joint US-UK contractor teams for evaluation.

On the basis of this work an SHF A/J Interoperable Waveform Standard has been agreed by the nations involved. This is being used in the prepartion of a NATO STANAG. The STANAG is being drafted by France, which has declared a desire to participate in the UM programme.

Bids for full-scale development of the UM are being invited, with a view to the first production models being available in 1995.

It is the declared intention of NATO to procure the UM for the HTT programme. However, there is concern as to whether it will be available in time for the earlier HTTs, in particular those for the SHAPE AWHQ.

2.4 FUTURE REQUIREMENTS

2.4.1 Need for Requirements

Even though the Group was not constrained by any need to have stated operational requirements it was nevertheless found necessary to postulate some generic repuirements so as to limit the scope of the studies to what may be relevant to NATO and to determine and try to correct the deficiencies of the existing systems.

2.4.2 The Future Role Of SATCOM in NATO

Plans for the evolution of NATO SATCOM during the lifetime of the NATO IV space segment indicate a period of consolidation for the inter-static network followed by a reconfiguration of the network as evolution towards the C3 Architecture gathers pace. At the same time there will be a proliferation of transportable SHF SATCOM terminals for crisis management, political consultation and emergency communications.

a) The Future of the Large Static SGTs

Although fibre-optic links will provide a viable alternative to SATCOM for long-haul common-user circuits in the long term, it seems unlikely that any large static SGTs will be closed down in the near future simply because alternative media have become more cost-effective. However, the need for a major and costly overhaul in order to extend the life of such an SGT may be reason enough to opt for closure. The original twelve SGTs were designed for a fifteen year life and were brought into service during the period 1970-1972. However, during the period 1982-1985 they were modified to bring them up to the standards required of the SATCOM III ground segment. This will have extended their life somewhat, though certain major subsystems such as the antenna were unaffected by the modification. The nine new SATCOM III terminals, brought into service between 1982 and 1985 also have a fifteen year design life and hence may be expected to require replacement or a major overhaul between 1997 and 2000. Thus closure of the inter-static SATCOM network in favour of terrestrial alternatives at the end of the NATO IV era appears to be a realictic option. However there is likely to be pressure to retain at least some of the the SGTs for political reasons, and also to anchor circuits from mobile (in particular, shipborne) terminals.

b) The Future of the HTTs

Deployment of the highly-transportable terminals (HTTs) as currently envisaged is unlikely to be completed much before the year 2000. On present plans, all HTTs will, by this time, be equipped with the Universal Modem which is intended to operate through wideband transparent transponders. To capitalise on the investment in the HTT subsystem, a future NATO space segment will, like its predecessors, need to provide wideband SHF transparent channels. Some of the roles of the HTTs, such as crisis management and political consultation, are so fundamental that their continuation beyond the year 2000 is scarcely in doubt. Others such as the support of AWHQ and TNF emergency communications are more dependent on current NATO military doctrine and therefore more likely to change in the future. However,

the flexibility of the HTTs will allow them to be transferred from one role to another at short notice. It is anticipated that following the year 2000 NATO will continue to find applications for all HTTs that have not exceeded their design lifetimes (and probably for some that have) so long as they can maintain an adequate data rate for the application concerned in the presence of the perceived ECM threat at that time. The design lifetime of the HTTs is likely to be at least fifteen years, so many of the HTTs could in principle remain operational beyond the year 2010.

c) Maritime SATCOM

The support of maritime forces has long been established as a major role of military SATCOM, both NATO and national. As more national ships liable to come under NATO command become equipped with SATCOM terminals the demand for capacity is likely to grow. If a future space segment serves both NATO and national needs then shipborne terminals of collaborating nations will need to be supported whether the ships are under NATO command or not.Maritime applications stress ECCM and LPI, factors which could prompt an early move to EHF. In the current NATO SATCOM system, large anchor stations are required to terminate the ship-shore and shore-ship links and to provide an interface with the common-user network. Consideration should be given to the replacement of these by smaller, mobile terminals.

A limitation of the present NATO SATCOM system in its maritime role is the lack of coverage beyond a latitude of about 70°N. Certain nations who rely on NATO SATCOM for national naval communications, or with whom NATO might wish to collaborate in the development of future joint NATO- national satellites, have indicated a potential requirement for polar coverage, as has SACLANT (see Appendix 2C). This suggests that alternatives to the geostationary orbit should be seriously considered for the post-2000 era.

National navies have peacetime SATCOM requirements in addition to their ECCM SATCOM requirements. At present these are met at UHF using US satellites of the FLTSATCOM system. In the next decade this capacity will be supplemented by the UHF channels on NATO IV and on the UK's Skynet 4 satellites. These UHF channels do to not have "military" characteristics and offer no ECCM protection. Furthermore, the capacity provided for a given payload mass is very small compared to SHF. It is therefore suggested that if this requirement for "peacetime" UHF SATCOM continues beyond 2000, serious consideration should be given to renting channels on other satellites-possibly civil- as this could well prove cost-effective.

Another potential requirement is for an ECCM link to submarines, essentially one-way since submarines would normally maintain radio silence, but possibly two-way using advanced LPI techniques (e.g. at EHF). This is a particular case where extended polar coverage is of interest. Ideally a link to a submarine should be capable of operating when the vessel is still submerged, and this is technically possible using a modulated blue-green laser beam (see section 8.6). However , the technology is still immature and it seems unlikely that this would be a viable option for NATO until well beyond 2000

d) Airborne SATCOM

A very likely requirement, new to NATO, for the post-NATO IV era will be for SATCOM links to aircraft. There is considerable national interest in the use of SATCOM to provide reliable beyond-line -of-sight links to aircraft: The UK is known to be considering use of SATCOM for Maritime Patrol and AEW aircraft, and is currently conducting trials with the Marconi MASTER terminal fitted to a demonstrator aircraft and using the VSC-330 frequency-hopping modem. The US has had an SHF SATCOM trials aircraft flying for several years and is known to be developing EHF airborne terminals as part of the MILSTAR programme. France is also understood to have a requirement for airborne SATCOM. An obvious application in NATO would be to the NAEW aircraft. The main difficulties with airborne SATCOM at present are the limitations on antenna size and transmitter power. In the case of a transparent SHF satellite channel, these make it necessary to anchor the link at a large SGT. On the SGT-aircraft link, the SGT has to capture a disproportianately large share of the satellite's EIRP. However, in the future it should be possible to avoid use of such a large anchor station by use of on-board processing. Also, use of conformal array antennas on the aircraft in place of steerable dishes should allow higher gains and will allow much more freedom in siting the antenna. If such arrays can be made active, higher aggregate transmit powers may be obtainable as well. Another way of obtaining higher EIRPs from airborne terminals is of course to implement the system at SHF.

e) Satellite-Borne Switch

A possible application of a future SATCOM system that may merit further study involves the use of on-board processing to provide an inter-network gateway function. Although the basic role of SATCOM in a NATO ISDN will be to provide point-to-point bearers, consideration could also be given to providing a switching function on board the satellite. This of course implies on-board data regeneration, but regeneration could in any case be a feature of an advanced ECCM SATCOM architecture. A satellite based switch linked to most if not all the terrestrial gateway nodes would significantly increase the robustness of the NATO ISDN and could well represent more efficient use of SATCOM resource than providing a set of point-to-point SATCOM bearers to supplement connectivity. It would e possible to re-load the switch software from the ground to cater for system upgrades etc.

2.5 REFERENCES

[2.1] "The C3CS Goal Architecture", STC Technical Memorandum TM-867 (Draft), 1989. (NATO Confidential)

[2.2] Report of the NATO SATCOM Enhancement Working Party, Aug. 1987, Ref. AC/317-(WG/1) WP/25. (NATO Secret)

[2.3] Payzin, A.E. "NATO IV Satellite Loading Plans", STC Technical Memorandum TN-174, Feb. 1986. (NATO Restricted)

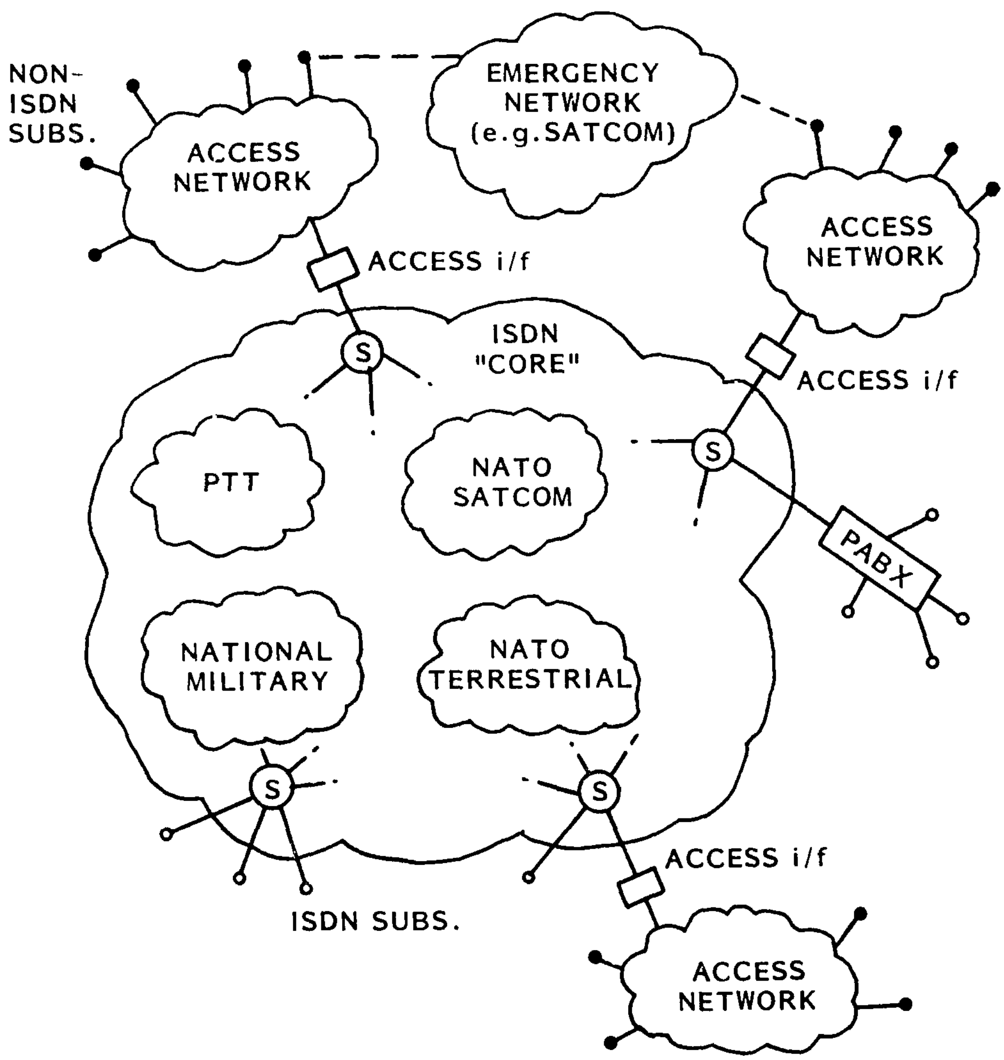


Fig.2.1 A NATO ISDN Concept

APPENDIX 2A

INTEGRATED SERVICES DIGITAL NETWORK (ISDN)

NATO decided in 1984 that most of the NATO terrestrial communications requirements would be met in the future by the strategic military communications networks that are today being designed and some being implemented by the Member Countries. All these networks largely follow the CCITT IDN/ISDN standards and recommendations and adopt the International Standards Organisation's (ISO) Open System Interconnection Reference Model (OSI/RM). These digital common-user grid networks provide mission related, situation oriented, low-delay "teleservices" such as plain/secure voice, facsimile and non-interactive and interactive data communications. These are enhanced by "supplementary services" such as priority and pre-emption, secure/non-secure line warning as well as closed-user groups, call forwarding and others. The switching subsystem supports three types of connection methodology namely, semi-switched connections, circuit-switched connections, and packet/message switched connections. The circuit switching technique use is the byte-oriented, sychronous, time-division-multiplexed (TDM) switching in accordance with CCITT standards. The basic channels are connected through the network as transparent and isochronous circuit of 64 kb/s or nx64 kb/s where n is typically 32.

The basic channel structure used in ISDN has T and S reference points and consists of two B channels at 64 kb/s and one D channel at 16 kb/s. One or both of the B channels may not be supported beyond the interface. The B channel is a pure digital facility (that is, it can be used as circuit-switched, packet-switched, or as o non-switched/nailed facility),while the D channel can be used for signalling, telemetry, and packet-switched data. The basic access allows the alternate or simultaneous use of a number of terminals. These terminals could deal with different services and could be of different types.

The primary rate B-channel structure is composed of 23 B or 30 B channnels depending on the national digital hierarchy primary rate, that is, 1544 or 2048 kb/s and one D channel at 64 kb/s. PABX connection to the T reference point can use (depending on its size) multiple basic channel structure accesses, a primary rate B-channel structure, or one more primary rate transmission systems with an common D channels. The primary rate H-channel interface structures are composed of Ho channels (384 kb/s) with or without a D channel, or an H1 channel (1536 kb/s). H channels can be used for high-fidelity sound, high-speed facsimile, high-speed data, and video. Primary rate mixed Ho and B-channel structures are also possible. Subrate channel structures are composed of less than 64 kb/s channels and are rate adapted and multiplexed into B channels.

Future evolution of the ISDN will likely include the switching of broadband services at bit rates greater than 64 kb/s, at the primary rate, as well as switching at bit rates lower than 64 kb/s which are made possible by the end-to-end digital connectivity. Table I shows some typical service requirements for civil and also for military applications.

In the ISDN environment, the use of common channal signalling networks significantly reduces the call setup and disconnect times; use of Digital Speech Interpolation (DSI) can enhance the transmission efficiency on a cost-driven basis.

Packet switching which allocates bandwith on a dynamic basis, has become the preferred technique for data communications. In addition to utilising the bandwith more efficiently packet switching permits protocol conversion, error control, and achieves fast response times needed for interactive communications.

Looking ahead into the future both for military and civil applications, we see good prospects for the integration of voice and data traffic. Investigation of different techniques permitting integration of voice and data traffic in one network has been a subject of ongoing research for more than a decade. These techniques include hybrid switching, burst switching, and packet switching for speech and data. A common objective of all these techniques is to improve efficiency of speech connections in comparison with the circuit-switched network, with minimal degradation to speech quality as a result of clipping and message delay.

Hybrid switching can achieve acceptable voice message delays. However, lower transmission efficiency and higher complexity than packet-switching concepts render it unattractive for application in public switched networks.

Burst switching achieves high transmission efficiencies and low voice message delalys. It is an attractive concept, but high costs associated with the development of a new family of swicthing systems and the lack of evolutionary migration paths for implimentation make it unsuitable for public networks.

The attraction of speech packet communications lies in the relative simplicity of packet-switching concepts, and the fact that computer systems for data packet switching can be adopted for speech packet communications. While existing protocols for packet data communications such as x.25 are not suitable for achieving small fixed delays necessary in speech packet communications, significant progress has been made in developing new protocols under the sponsorship of the Defence Advanced Projects Agency (DARPA) and the Defence Communication Agency (DCA). While still in a developmental stage, speech packetisation increasingly appears to be the prime contender for future voice/data integration in common-user networks.

Another speculative impetus for speech packet communications lies in the potential for voice recognition and direct speech input to program, command, and control the operation of artificial intelligence machines. Speech packet communications are ideally suited for such applications.

Table 1
Some service requirements

Service	Bandwith Requirement	ISDN Channel Type B	ISDN Channel Type D	Facilities: Circuit Switched	Facilities: Packed Switched	Facilities: Channel Switched	Facilities: Overlay
Telephone	8,16,32,64 kb/s	X		X			
Interactive Data Communications	4.8 - 64 kb/s	X	X		X		
Electronic Mail	4.8 - 64 kb/s	X			X		
Bulk Data Transfer	4.8 - 64 kb/s	X		X			
Facsimile/ Graphics	4.8 -64 kb/s	X		X			
Slow Scan / Freeze Frame TV	56 - 64 kb/s	X		X			
Compressed Video Conference	1.5 - 2 Mb/s (Primary rate)					X	X

APPENDIX 2B

NETWORK ASPECT

Aspect of SATCOM with reference to the different roles that SATCOM can play as a part of a fully integrated switched network are:

- -Transmission performance
- -Signalling and routing
- -Network surveillence and control

and -Transition to future systems.

The following points can be made with respect of the above:

i) SATCOM links can be designed to meet any transmission standards with respect to error-rate and slip (provision of elastic buffers synchronised to the network time).

ii) Since satellite links can provide a "node-skipping" capability they give an alternative in many cases to an all-terrestrial route within the switched network. The routing system should therefore be such taht the use of SATCOM capacity should be maximized (high-usage mode) without however causing unnecessary pre-emptions (search for free terrestrial paths before SATCOM circuits are pre-empted).

iii) Problems of signalling over satellite links arise because of

a) the long propagation delay
b) the need for a given earth station (ES) to serve two or more switching nodes
c) the numbers of channels per link over the satellite will be different from the numbers of channels per link over the tail circuits
d) the requirement to operate with reduced channel capacities in the case of jamming

If survivability is not a problem, a non-associated technique could be considered against (a), with which the signals concerning calls over the satellite links could be carried over the signalling channels of a number of terrestrial links in tandem. The associated common-channel technique for SATCOM internodal links can be used without requiring more than one signalling channel on the link between a pair of earth stations (ES). This can carry the signals generated by several internodal links. To allow the nodes using one ES to share a single satellite signalling channel to the nodes using the distant ES, a signal multiplexing arrangement would be required at each ES.

Signalling channels could be set up directly between nodal switches without requiring a signal processor at the ES, but at the cost of additional satellite capacity for the separated signalling channels.

iv) The satellite signalling system would also be called upon to deal with degraged operation under jamming conditions when only the AJ protected links may be operative. Some or all of the speech channels on any one internodal satellite link would be disconnected as determined by the Network Control Authority. The switch must know which circuits are still available for traffic and the signalling system must cater for this.

v) In cases where access links connecting users to the nodal switches are provided via SATCOM using, for instances, low-cost or transportable ES's, the node terminating such links should be one which is near an ES. To satisfy the transmission requirement for not more than one satellite hop, the onward routing of calls by the parent nodes would be wholly terrestrial in such cases. When such a routing cannot be found the satellite-connected parent node would cause calls to be diverted to the terrestrially-connected parent nodes, since the access switch would also have terrestrial access links. The alternative, of allowing a direct satellite link to be established between access switches, would add functions (e.g., routing and control of COMSEC equipment) to the access switches since no node would be available to perform them, as well as demanding additional satellite capacity. The use of demand-assignment technique to establish such direct links would ease the latter constraint but would require further functions (in addition to routing, etc.) to be performed by the access switches.

vi) The question of transition from an existing system to a series of future systems is an important issue which should be considered in designing new architectures. It is, indeed, one of the essential virtues of the architectures being proposed in this report (see Chapters 10 and 11) that this aspect of easy transitioning is inherent in that the SHF assets are catered for as well as well the EHF in a changing ratio over the time frame in question. Moreover, the satellites being postulated in the report, though use on-board processing, are also useable transparently.

APPENDIX 2C

NATO INPUT FOR WG-13

NORTH ATLANTIC MILITARY COMMITTEE
COMITE MILITAIRE DE L'ATLANTIQUE NORD

NATO UNCLASSIFIED

15 May 1986

MEMORANDUM FOR THE AGARD AVIONICS PANEL EXECUTIVE
(Attn: Lt.Col. Stratton)

SUBJECT: Satellite Communications Working Group

Reference: AGARD letter AVP/31, 2 April 1986

1. The IMS has reviewed the Terms of Reference (TOR) for the AGARD Avionics Panel (AVP) per subject request and provides the following comments for consideration:

   a. In the light of recent NATO SATCOM IV directions, it would appear that the proposed study would provide long term guidance only. It should therefore address the broadest possible spectrum of military users. Consideration could be given to including under STATUS: D), maritime communications, including submarines, blue-green lasers and communications in the polar zones.

   b. Reference "STATUS: D) Propagation effects" - area should include nuclear/nuclear scintillation effects if it does not already.

2. If established, it is requested that the AVP Working Group establish close liaison with appropriate NATO agencies/committees e.g. STC, NACISA, Satellite Working Group. These groups have recently addressed several of the proposed working group objectives in designing NATO SATCOM IV.

3. This letter confirms data provided by telecon on 15 May 1986 to support AVP business meeting discussions.

Russell C Morrison Jr

R.C. MORRISON
Lt Col USAF

Note: Prof. Ince had a discussion, at the AVP 1989 fall symposium held at Colorado Springs United States, with Admiral Mason of SACLANT concerning military requirements. He confirmed that communications in the polar region for aircraft and also communications with submerged submarines are of great importance to SACLANT

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

NORTH ATLANTIC TREATY ORGANIZATION

AVP/31

LT COL R. MORRISON
IMS, ASI DIVISION
NATO-OTAN
1110 BRUSSELS BELGIQUE

2 APRIL 86

SUBJECT: Satellite Communications Working Group

In March, 1986, the National Delegates Board of AGARD approved an Avionics Panel proposal to establish a Working Group on Satellite Communications to be headed by Professor Doctor A.N. Ince of Turkey. The proposal, as well as Dr. Ince's address and telephone/telex numbers, are attached as Inclosure 1.

This Working Group and the Avionics Panel request input from NATO on Satellite Communications problems. They welcome guidance from NATO on areas of particular interest to NATO on this subject.

The next Avionics Panel business meeting will occur on 16 May 1986, in Bolkesjoe, Norway, where this Working Group will be discussed. Should you in NATO have any suggestions as to areas of interest not covered in the proposal, or areas that require special emphasis, we would welcome your input before this date so that it can be brought to the attention of the entire panel.

Michael V Stratton

MICHAEL V. STRATTON
LT COLONEL, US ARMY
EXECUTIVE, AVIONICS PANEL

1 ~~INCL~~ W/D
~~as~~

cf: DR. G. HUNT, Avionics Panel Chairman
DR. A.N. INCE, Director, WG 13

CHAPTER 3

COMPARISON OF NATO AND OTHER SYSTEMS

3.1 OBJECTIVE

The purpose of this chapter is to identify and analyse both the similarities and the differences between NATO satellite communications systems and other civilian or military systems.

Two aspects will be considered separately:

- -the type of services to be provided
- -the environment in which these services have to be provided.

This analysis attemps to contribute to the identification of the fields in which NATO could draw some benefit from the R and D carried out for other civilian or military applications and those where specific R and D actions should be undertaken for the sole benefit of NATO.

3.2 SERVICES AND SYSTEM CHARACTERISTICS

3.2.1 Civilian SATCOM Systems

Among the various existing or future civilian systems, the following categories can be defined.

<u>3.2.1.1 Worldwide High Capacity Point-to-Point Services</u>

This category essentially refers to the Intelsat system which provides high capacity intercontinental trunk connections through geostationary satellites.

The world traffic is divided between cables and satellites and the general policy consists of keeping some balance between these two communications media.

The evolution of the Intelsat system is oriented towards more flexibility and adaptability to the customer needs.

A certain amount of reconfigurability has been introduced in the antenna subsystems. An on board microwave switching matrix is being tested on Intelsat VI. Future trends are towards on-board baseband switching and more use of Ka band to reduce congestion of geostationary orbit.

Intersatellite links are being considered in order to reduce transit time in links using more than one satellite. EHF or optics can be used for these links.

<u>3.2.1.2. Regional or National Point-to-Point SATCOM Systems Using Geostationary Satellites.</u>

A number of such systems are now in operation either using specific satellites (Brazil, Mexico, various US domestic systems, Indonesia, Australia, Eutelsat, Arabsat, France, Germany and in the near future Italy, Spain , Turkey, etc...) or using leased Intelsat transponders such as Algeria, Zaire and others).

Each of these systems provides a variety of numbers of point-to-point connections. The number of stations may vary from a dozen to several hundreds. The mode of access is either SCPC in the case of many small stations or FDMA (at base band level) or TDMA for larger capacity links.

During the last ten years, the trend has been to reduce the dimensions of the ground antennas, making them accessible to more users. One of the limitations to this trend is obviously the frequency congestion of the geostationary orbits which imposes limitations on the beamwidth of the ground antennas.

In addition to bilateral voice and data transmission, the concept of data broadcast has been introducced by various curtomers and this will result in an important development of the V-SATs (very small aperture terminals).

Due to frequency congestion, these systems (with the possible exception of those covering tropical areas) will more and more use the upper SHF (Ku band and Ka band).

<u>3.2.1.3. Regional or National Point-to-Point SATCOM Systems Using High Inclination Orbits</u>

These systems are typical of those implemented in the Soviet Union. The advantage for this country is obvious because such systems provide a good coverage at high latitudes. In addition, the eccentricity of the orbits (e.g. Molnyia) provides a relatively small angular motion of each satellite during several hours thus simplifying the tracking problem or making it even unnecessary for small aperture ground antennas.

<u>3.2.1.4. SATCOM Systems for Communications with Mobiles</u>

The only system currently in operation is Inmarsat which provides ship-to-shore communications. Since there is no real requirement for communications at high latitudes, this system uses geostationary satellites. The ship-to-satellites links use L-band and the shore-to-satellite links use C-band. Inmarsat is planning to extend its service to communications with aircraft.

Capability of mobile-to-mobile single hop links has not yet been implemented. This capability is being considered in future systems.

Several schemes have been proposed for communications with ground mobiles. The US-Canadian system M.SAT will use geostationary satellites.

Use of highly inclined orbits have been proposed to increase the elevation angle of the satellites at medium latitudes thus reducing the effect of obtructions due to mountains or buildings. This is the case for Sycomores and for Loopus which are described in chapter 10 of this report.

The main purpose of these future communication systems for mobiles is to achieve narrow band (voice or low speed data) communications between mobiles and the main communication infrastructure via a satellite and fixed communication centers.

The next requirement to come up will be to achieve single hop voice communications via satellite between mobile users. This will require on-board switching in the satellite first by command from control centers(2d generation) and possibly later on by automatic switching without action of the ground control centre except for network reconfiguration (3d generation).

<u>3.2.1.5. Data Relay Systems Between Low Orbit Satellites and Ground Stations.</u>

Presently one such system is in operation: the TDRS system of the United States. Another similar system (DRS) is planned to be implemented by ESA during the next decade.

In both cases the purpose of the system is to provide communication between low orbiting satellites (including manned vehicles) and the ground via geostationnary satellites thus reducing the need for a large number of ground stations exchanging data with these low orbiting satellites.

3.2.2. Military SATCOM systems

Three countries have implemented military SATCOM systems in the western world. In the chronological order, they are the USA, the United Kingdown and France. The main requirements for such systems can be summarized as follows:

- on the national territory, to complete and to provide a backup to the most critical terrestrial links which may belong to civilian networks or to specific military networks,

- outside the national territory, to provide an independant communication system with military units located outside the national territory,

- to provide communications with mobiles beyond the usual range of ground communication facilities (terrestrial vehicles, ships, aircraft).

Therefore, these systems must provide to the military users types of services which are similar to most of those described above in the various categories of civilian systems.

- At least the USA need worldwide trunk connections but certainly with capacities smaller than those needed by Intelsat.

- Any military satellite must have a wide coverage plus some spot beams, steerable or not, to provide better performance over specific areas.

- The ground segment of any military satellite communication system must include not only fixed stations but also, transportable and mobile ground terminals. Part of these requirements are therefore similar to those of a civilian mobile SATCOM system except that more complexity is acceptable for the military mobile terminals. They need to be more protected than those designed for commercial use.

- Military requirements may result in the need for communications with high latitude (polar) region. Existing systems in the western world do not satisfy these requirements and high inclination orbits will have to be considered in the future.

- Military SATCOM may contribute to improve communications with submarines.

To summarize the above, one can say that the services to be provided by military systems represent a synthesis of most of the services to be provided by the various civilian networks:

- trunk communications
- point-to-point communications covering both hemispheres
- communications with mobile users
- communications requiring high inclination orbits (although the justification for these orbits may not be strictly the same as for civilian systems).
- communications with submarines are specific to military requirements.

3.2.3. Peculiarities of NATO SATCOM

The services to be provided by a NATO SATCOM are generally the same as those of any military system. However the multinational nature of NATO and the existence of national military SATCOM's somewhat alters some of the requirements:

- The NATO countries cover nearly half of the northern hemisphere. The existence of several regional commands of NATO spread over this area justifies the need for permanent trunk connexions via satellite which, in case of war, may constitute a full backup to more vulnerable terrestrial links.

- A NATO SATCOM can also provide communications between the military units made available to NATO by the various countries, whether these countries have their own SATCOM or not. On that aspect, the characteristics required from a NATO system do not differ from those of a national SATCOM systems.

- The coexistence of national systems and of a NATO system within NATO makes it necessary to ensure interoperability between national and NATO systems in order that each system can provide a back up to the other at least for some specific links.

- NATO countries having their own SATCOM may use it for tactical communications (i.e. mostly communications with small transportable or mobile terminals) but one should keep in mind that a NATO SATCOM should provide some capacity of tactical communications for three reasons:

  - providing some of this capacity to military units of nations who do not have their own SATCOM

  - Provide some backup to nations who have their own SATCOM's

  - provide in any case a possibility of access to small terminals as these may be needed in case of emergency for non tactical communications.

- There is definitely a need for a NATO SATCOM to provide communications with polar regions and the solutions to the problem described above for civilian systems are similar to those which could be applicable to future NATO SATCOM's ac well as to national military SATCOM'.

- Although this is rather specific to tactical needs (e.g. mobile or transportable station operating in mountain areas) the solutions considered for civilian systems of communications with mobiles such as the use of high inclination orbits are applicable to NATO or to national SATCOM systems.

- Communications with submarines and protected communications constitute a requirement for NATO SACTOM's as well as for national military SATCOM's.

As a general conclusion of this paragraph it can be said that a NATO SATCOM and military SATCOM's in general, must provide a wide choice of services similar to those provided by the various present and future civilian systems and that the differences between national and NATO military SATCOM's are the relative proportions of the various categories of services. Emphasis must be put on the need for interoperability of the NATO and the national SATCOM's.

3.3 ENVIRONMENTAL FACTORS

As a general rule, the environment for which a military and of course also a NATO SATCOM system must be designed is much more hostile than the environment of o civilion system.

Several categories of environment can be distinguished.

3.3.1. User Environment

What is called here users environment is the status, at any given time, of the need for communications expressed by the users of the system.

In the civilian systems, at least those for which statistics have been established for a long time, it is relatively easy to define the

capacity required to satisfy the users . The daily and the seasonal variations of the traffic are well known and a good forecast can be made with regard to their evolution. Unexpected variations of this evolution are slow enough to give time to correct the situation first by using the in built capability of the satellites, for reconfiguring their coverage or their transponder layout and for larger variations, by introducing spare satellites stored in orbit or on the ground.

The variations of the traffic military SATCOM's are much less predictable and they may occur much faster. Four different environmental conditions can be conzidered with regard to users requirements

- peace time
- manoeuver period
- crisis time
- war time

The first two conditions are relatively easy to forecast. Analysis of the traffic recorded during the second one (manoeuver) may contribute to a limited extent to evaluate or to check tle evaluations of what is likely to happen in crisis time or in qar time.

In these last tqo conditions, the variations of the traffic can be very high, surges can occur in certain areas within a few days, the area where the need and/or the capacity of traffic occurs may change quickly in a matter of days.

Therefore a military (and NATO) SATCOM system must be designed with an overcapacity as compared to the "peace time" raquirements and the transponder and antenna subsystems of the satellite must have some built in capacity of being reconfigured to meat daily changes of the traffic requirements

3.3.2 Enemy Environment

This environment determines the major differences between civilian and military systems.

A civilian SATCOM must survive the threat imposed by its operations in space i.e., thermal variations, eclipse conditions, natural radiations.

Some degree of protection is used in civilian SATCOM's to make the communication more secure but, for the time being, it is left up to the customers whether they wish to encypher their messages, whether they are transmitted via satellite or not.

The command signals sent to the satellite are in certain cases encyphered to protect the satellite against any unwanted (deliberate or accidental) action.

All the above considerations must be taken into account in the case of a military SATCOM but for these military SATCOM (including NATO) the threat originating in potential enemy actions is governing.

This threat also exists for civilian SATCOM's which may carry military circuits but economic considerations make it impossible to afford the cost of countermeasures or any other type of protection in civilian systems.

A large part of the report being devoted to the threat, only the main headlines are recalled here in order to mephasize the porticular conditions that a military (and NATO) SATCOM system must be designed for:

- jamming and interception of the communications
- jamming or take over of the command link
- nuclear effects on:
- satellite
- propagation
- ground segment
- attacks against the ground segment (including sabotage)
- attacks against the satellite:
- radiation
- physical attack

In this field there is no difference between the solutions to be applied to a NATO SATCOM as compared to a national military SATCOM. Some techniques also applicable for civilian SATCOM's may contribute to solve the military problems (laser communications, antenna arrays, on board switching and processing, superconductivity,...) but the complete solutions to these problems including the necessary R and R are specific to the military domain. The relevant subjects of R&D are the same for national SATCOM'S and for a NATO SATCOM system

CHAPTER 4

ISSUES AND FUTURE TRENDS IN SATELLITE COMMUNICATIONS

4.1 OBJECTIVE

At the very outset the Group found it useful, even necessary, to list all the issues and future trends in satellite communications that could be delineated for debate with a view to using the answers (including doubtful comments and no answers, both justifying further work) produced in the development of a detailed work program to be carried out by the Group in line with its Terms of Reference.

4.2 ISSUES

The following are the issues which have been formulated for debate:

i) a) What differences exist, now and in the future, between civil and military SATCOM systems?
b) Can one hope of using civil systems for NATO purposes?
c) Can one rely on civil developments only (e.g. NASA, ESA, Intelsat etc.)for military use?
d) How different is NATO SATCOM from other international/national systems (Intelsat, MILSTAR, EUTELSAT, EUTELSAT, SKYNET)

ii) Can this WG do something new, original and helpful to NATO and The Nations that the other existing bodies cannot?

iii) What are the national programs with relevance to future NATO SATCOM (SDI, NASA, ESA, SKYNET,...)? Can we have access to the results of their developments and forecasts?

iv) What is the expected impact of SDI? What possibilities exist for technology transfer?

v) What are the attributes of an ideal SATCOM system?

vi) Do we expect revolutions in the next 30-40 years or evolutionary developments in SATCOM?

vii) What will be the impact of ISDN on SATCOM development and the role of SATCOM in military ISDN's?

viii) What role would one foresee for SATCOM in the years beyond 2000; strategic, tactical, special purpose, multipurpose, army,navy, submarine, air?

ix) Are there new and imaginative system architectures in sight?

x) How are advancing technologies going to affect threat to SATCOM and the countermeasures?

xi) Are there hazards to SATCOM which we do not have today but foresee for the future?

xii) What would be the limits on weight, size, control on-board power and processing, price-performance ratio, launch capability?

xiii) What would be the consequences of orbital and frequency crowding?

xiv) What are the frequency bands to consider for future systems?

xv) What are the expected developments in optoelectronics?

xvi) Can we develop future threat scenarios based on our technology forecast?

xvii) What are the relative advantages of geosynchronous, supersynchronous, and sub-geosychronous communication satellites?

xviii) What are the relative advantages/disadvantages of a) very large satellites in synchronous orbits with small and cheap terminals (is there a point of diminishing returns with regard to economies of scale?) and b) several low-orbit satellites using packet switching techniques?

xix) What are the relative advantages of single-dual-or multi-purpose satellites (all communications transponders working at different frequencies etc. or different applications on one platform?

xx) What is an optimum satellite lifetime? What about servicing and replenishing spacecraft in orbit?

xxi) What are the considerations for optimising the ground segment or the space segment?

xxii) In optimising spacecraft design, what is best to maximise: power, antenna, on-board processing, stabilisation system? All or none of the above?

xxiii) Is there an optimum point of reliability and complexity with respect to on-board redundancy, on-ground or in-orbit spares etc.?

4.3 RESPONSES

The following responses were given by the Group to the issues and questions posed in Section 4.2 above. The paragraph numbers below refer to those in Section 4.2. It is to be noted here that some of the views given below were somewhat modified during the course of the studies as reported in the later Chapters of this report.

i)(a) -The technologies used for both the civil and military SATCOM are basically the same. Military systems are currently hardened to varying degrees against ECM, radition effects (including nuclear burst), EMP, and laser attack. This will also be the case for future systems. The degree of hardening will depend on threat level and budget considerations.

-Military customers will always be prepared to pay some premium in cost or data rate to obtain:

Network flexibility
Earth terminal size for portability and reduced visibility.
Security against eavesdropping, jamming and comint.

i)(b&c) NATO cannnot rely solely on civil systems, but civil system should be used as part of a media mix.

The Defence Communications Agency policy has been as follows:

33 % Landline
33 % Civil SATCOM
33 % Military SATCOM
33 % Line of Sight
giving 33 % Spare capacity

A similar policy would be both necessary and prudent for future systems. Civil systems will have improved hardening characteristics arising out of new technology. For example radiation hard semi-conductors are already in use to prolong orbit life and are expected to improve further distributed transmitters and receivers associated with phased array systems will be more resistant to laser attack simply on an area basis. Civil systems will also use on board-processing which will give some protection against jamming.

The greatest advantage of using civil systems for some of the military requirement is the increase in network routes which can be available for critical communications.(*)

i)(d) NATO SATCOM is different from comparable civil systems since it relies on a media mix as outlined in i(b&c) above. It is also different from comparable military systems because of the NICS architecture which calls for a limited range of services(**), and NATO has no organic Threat Assessment capability. Skynet 4 satisfied both the UK national requirement and that of NATO: MILSTAR is primarily a hardened EHF, low traffic capacity satellite for critical communications only.

ii) The Group can put forward new ideas which could, however, be speculative. It can also collate technical ideas from a range of sources which , if integrated, may provide a new approach or concept. The collation of ideas can particularly be helpful to NATO if NATO explores further these ideas by studies in NATO and/or within the Nations.

iii) The primary national programs include the following: DSCS, MILSTAR, FLEET SATCOM, SKYNET 4, SKYNET 5 studies, SYRACUSE, and the military SATCOM plans for Germany, Italy and the Netherlands and Canada. Although not in NATO it may be useful to establish what Australian intentions are for Milsats as they have strong connections with UK/USA thinking but often an indepentent view.Canada is known to be developing a prototype processing payload called (FASSET) for a possible future military SATCOM. WG-13 members will try to collect information on these national systems and if this is not posssible we ask NATO and the Nations concerned.

iv) A view held by some members is that NATO access to SDI or SDI technology is unlikely at a sufficiently detailed level to be WG-13. This however has not been tested

v) The attributes for an ideal military SATCOM system are easy to define in general terms viz:

a) Complement terrestrial transmission systems, including fibre optic links with minimum impact on switched network architecture and communications to mobile terminals.

b) Be reliable and highly available

c) Have a high ECCM performance via nulling antennas and spread spectrum techniques.

d) Be largely autonomous.

e) Capable of wide area coverage, not only via earth coverage antennas but also by beam steering and on-board switching.

f) Be hardened against attack by laser, fragmentation (Kinetic Energy) weapons and RF, high energy.

g) Have high transmission quality ie., low BER.

h) Be affordable.

Much more difficult, is to independeenly identify values for each of these attributes.

vi) Yes and no. For the space sector satellite autonomy will revolutionise traffic resource allocation and reduce the ground satellite control costs.

For the ground segment the massive investment in military SATCOM ground equipment suggests that no revolutionary changes in network architecture will be possible within NATO budget. Satellite switched TDMA may however provide sufficient advantages in performance in some future military scenarios to justify a new system to complement the current SSMA/FDMA mix.

Growth is most likely in the areas of submarine, and mobile communications to aircraft and special purpose forces. The following may be regarded as "revolutionary technologies":

- superconductors for analog processing elements
- digital signal processing
- AJ algorithms programmable in space
- use of MMIC in large scale
- Robotics for repair and replacement
- autonomous operation
- use of novel technologies for submerged submarines (lasers, ELF antennas in space)
- use of SAW and acousto-optical technology (for beam forming and spectral anlysis).

vii) Force standardised protocols onto satillite switches and perhaps lower propagational delay.

viii) Multirole with emphasis on small mobile terminals

ix) Perhaps- but not evident.

x) A move to EHF and fast frequency hopping spread spectrum methods will provide greater performance margins against jamming which will be necessary against compact jammers using such devices as gyrotrons. There is no permanent advantage achievable by either satellite or jammer conceivable over an individual satellite life time

xi) Yes. High power lasers and Kinetic energy weapons.

xii) Taken separately or together-absolute cost.

---

*It is understood that INMARSAT cannot be used for military mobiles

** The NATO system will have some dynamic response capability to changing threat level. (MOU, traffic reduction, nulling, more power to SSMA carries.)

xiii) a) Orbital Crowding

Some or all of the following could be employed.

- Very large spacecraft at allocated positions. These could start small and grow by robotic assemly in orbit. EM interference could well become a limiting factor in size.

- Clusters of spacecraft (possibly linked by tethers). Separated spatially to reduce the EMC problem, but with the centre of gravity at the allocated positions individual spacecraft in a tethered cluster could be positioned along an earth radius and still essentially remain in geostationarly orbit.

- Non-geostationary orbits such as Molnya or Loopus orbits would become more popular. This would however give rise to attendant complexity on the ground.

b) Frequency Crowding

The first approach would be to utilise to the maximum extent, the frequencies and bandwidths allocated by WARC to the military. Some re-use of frequencies are possible where the operational theatres are widely separated spatially. Re-use of downlink frequencies would also be possible at say EHF by the use of very precise antenna footprints for localised coverage areas.

Another more complex approach would be to use adaptive nulling both on-board the spacecraft and ground terminals. Antenna pointing requirements would then become more difficult to meet.

xiv) EHF in the 20-44 GHz. intersatellite links could use still higher microwave frequencies or lasers.

xv) *Uses for communications are limited in the lower* atmosphere by attenuation cloud cover particularly on the NATO operational theatres make optical links unreliable. Intersatellite laser links useful for large *volume data links (higher laser powers needed for* intersatellite links; new lasers and optical components for submarine communications).

Optical computers theoretically could provide massive parallel computing capability for on-board processsing, but has not developed at the same rate as the competition, ie the transputer. Optical systems so far have proved very massive and required mechanical stabilities and precision very difficult to achieve and maintain.

xvi) To some extent, witihin our timescale 2000-2030 some 40 years from now and if technological innovation continues at the present rate, any forecast in this direction will only be valid for perhaps 10 years at most.

xvii) The relative advantages/disadvantages are given below:

GEO : Tracking/Command easy.
Constant availability over a given area.
SUPER GEO : **Less susceptible to laser and fragmentation devices.**
One satellite provides intermittent, worldwide coverage.
SUB GEO : **Intermittent, worldwide coverage.**
Lower delay,loss.

More payload possible for a given launcher

The question should be rephrased to include various inclined orbits and compare those also to geostationary orbits.

xviii) *This topic cannot be readily addressed unless criteria* are set for availability and survivalibilty. A cost model can be constructed to address the relative economic advantages of several smaller satellites against very large satellites. A system called MEWS (see Chapters 10, and 11) can accommodate high power satellite transmitters as well as the very large satellite and consequently make the use of small and cheap terminals just as possible.

As the traffic requirement increases a point will be reached where satellite will be so large that in-orbit assembly will be necessary. This represents a step-function in costs which arise from multiple launchers, the addition of either robotics for assembly or man intervention.

Given an availability/survival criteria a trade off between hardening against physical attack in each case could be carried out. Several smaller spacecraft have an advantage in not having all ones eggs in the same basket and lower mass spacecraft are more manoeuvrable (ie less inertia) than larger satellites in avoiding physical attack. They also represent smaller target cross-sections.

Intuitively, several smaller satellites should provide the better system.

xix) The most obvious difference between multirole and single role satellites is that for several roles to be executed more satellites will be required by the single or dual than for the multirole case. *To achieve the* most benefit in availability/survival the larger number of satellites would of necessity also be multirole. If a *particular role was more critical, more hardening, more* ECCM capability could be built into that single satellite without compromising all. MILSTAR is perhaps an example of this approach. Launch costs for a satellite *represent about 30 % of the total cost to orbit. Even if* all satellites were launched piggy back by the same launcher the overall mass for the multiple case would be greater than the single large satellite.

As launch costs are usually expensive in dollars/kilo launched, a cost advantage would lie with the large multirole satellite. Again the trade off lies in how much value is there in greater availability and survival to justify a larger system cost. The current view taken by NATO, UK, France and Italy is that the multirole satellite is the only one affordable. The USA takes a different view, conditioned in part by its assumed global role.

xx) The life time of a satellite is in the main set by the following:

a) Degradation of solar cells.
b) Number of battery charge/discharge cycles.
c) The amount of fuel capacity for orbital station-keeping, re-positioning and de-orbiting.
d) The failure rates of the electronic and mechanical parts used.
e) The extent of redundancy employed.
f) The orbit used.

For low earth orbits (LEO), the degradation of solar cells is higher than for GEO due to higher levels of

radiation in the Van Allen belts. This also applies to inclined and highly elliptical orbits but to a lesser degree as the spacecraft spends less time in the low earth part of the orbit. For inclined and polar orbits batteries are charged and discharged by about 2.500 times more often than for GEO due to the number of eclipses seen-about 14 per day. In addition the spacecraft inertial attitude must be continuously changed to point the communication antennas perpendicular to the earths surface. This obviously requires more fuel.

Currently geo-stationary spacecraft have design lives of 7-10 years while inclined and low earth orbiters have between 4 and 6 years. In general it is possible to design to meet almost any lifetime, by carrying larger solar arrays more batteries and more redundancy. This approach is used to some extent but is limited by available budget and available launch platform to carry the increased mass. It is also determined by the overall system life time which for military communications is largely determined by the ground sector.

Major changes in system architecture which demand space sector changes are unlikely to occur more often than every 15-20 years. A possible driver for space sector change is increase in traffic demands. This has certainly taken place in both NATO and the UK, and has led to new spacecraft design. This has also been true for threat change. The optimum life time is therefore set by system requirements, available budget and not by available technology.

Servicing and refuelling in orbit is feasible at both LEO and GEO, but the space segment has to be designed and configured to enable this to take place economically. So far this has been carried out by the simple replacement of a fuel spent or faulty spacecraft. To minimise any down time an in orbit spare quiescent satellite has been used to replace the defective spacecraft. Several spacecraft have also been provisioned so as to statistically assure full system services.

Servicing by man intervention or by robotics at orbit replacement unit (ORU) level has been proposed for Space Station and Polar Orbiter for Earth Observation. At the altitude of inclined or polar orbiters man intervention is hazardous due to the high radition levels. Repair by robotics demands an onerous degree of standardisation of ORUs which increases mass and cost. At the ORU level, launch costs become comparable to replacement at full satellite level and no guarantee of continuity of service is provided as launch capacity is limited and not available on short term demand. Replenishment of fuel could be accomplished more easily by robotics. The replenishment point is determinable with some accuracy and appropriate launches organised in advance. The MEWS system incorporates this procedure.

xxi) This must be based on the minimum military requirement, the existing inventory and currently available technology. Assuming the requirement is met, no obvious fundamental parameter deserves optimisation in preference to others, except cost.

xxii) No one aspect deserves maximising unless a specific weighting in the requirement demands it. Optimisation of the spacecraft design will inevitably involve all features not just those listed in the questions.

xxiii) Probably. As reliability and complexity increases, so also does cost. There is probably a point where a large number of very small, single role and hence cheap satellites will provide a more survivable available system. Where this is in demand a study in depth against a well defined range of requirement scenarios.

Storage of space systems on the ground is less useful than in orbit.

a) It is readily available due to launch constraints.

b) The space environment is more benign than even cold storage on the ground.

c) Launch failures will have been experienced and maximum time made available for replacement satellite and new launch slot.

4.4 DETAILED WORK AREAS TO BE STUDIED BY WG-13

4.4.1 Services to be Provided by and Attributes of Future SATCOMS

a) Communications for fixed mobile and submerged platforms in all areas of interest to NATO including polar regions.
b) Attributes including AJ survivability covert etc.

4.4.2 Technology Assessment

a) RF and optical devices
b) Antennas
c) Switching and signalling
d) Processing signal bandspreading etc information data
e) Transportation(launch vehicles etc.)
f) Materials (light weight structures)
g) Support systems (power TTC etc.)
h) Software, Artifical intelligence and expert systems robotics
i) Superconductivity
j) Adaptive control systems
k) Advanced control systems
l) High RF power

4.4.3 System Architectures

Classification of architectures distributed and centralized
- Choice of orbits
- Number of satellites
- Protocols
- Life
- Vulnerability

4.4.4 Threat

a) Physical(space mines particle and kinetic energy weapon,laser and RF weapons burn out)
b) Electronic Warfare
c) Nuclear(EMP,radition)
d) Piracy
e) Signal intelligence

4.4.5 Countermeasures

a) Hardening
b) Redundancy
c) Maneuverability
d) ECCM
f) Low propatility of Intercept (LPI)

4.4.6 Propagation

a) mm wave propagation
b) Laser beam transmission in atmosphere and water

c) scintillation effects

4.4.7 Network and Transmission Issues (Implementation)

a) Transmission standards
b) Delay (important for SATCOM integration with terrestrial networks)
c) Connectivity
d) Signalling
e) Routing
f) Access
g) Synchronization

4.4.8 Spacecraft, Payload and Launch

a) Antennas
b) Front end
c) Processing
d) Transmitter
e) Filtering
f) Cryogenics
g) Reliability (redundancy, servicing and repair in space)
h) BUS

4.4.9 System Surveillance and Control (Management)

a) Space segment
b) Ground regment
c) Traffic
d) Autonomy issue
e) Repair, servicing and replacement

4.4.10 Ground Segment Design and Integration

a) Ground static and mobile terminals
b) Ship and submarine-borne terminals
c) Air-borne terminals
d) "Wrist-watch" and "brief-case" terminals
e) Integration with terrestrial switched networks (ISDN)

4.4.11 Space Transportation and Launch Vehicles

a) New launch vehicles
b) SDI and space stations
c) Orbital Transfer Vehicles and robots

4.4.12 Conclusion

It was ascertained that there was adequate collective expertise and experience in the Group membership to undertake the necessary studies in all the above areas as called for by the Terms of Reference.

CHAPTER 5

THREAT TO NATO COMMUNICATIONS

5.1 THE THREAT TO NATO SATCOM POST-2000

It is essential for NATO to have a clear perception of the threat to future SATCOM system at an early stage, since this will have a fundamental influence on the system design and cost. However, a basic difficulty in determining a threat level against which to design a future system is in deciding what is realistic as opposed to what is technically feasible. For instance, it can be argued that an enemy would only be prepared to invest in a particular type of facility (say an uplink jammer) up to the level where an alternative form of attack (e.g. physical destruction of the satellite) appeared more economical. Alternatively, an enemy might use the cost-to-deploy versus cost-to-negate ratio as a measure of viability: There would be no point in building, at enormous cost, an uplink jammer at the limit of technical feaasibility if it could readily be restroyed by conventional weaponry at the onset of hostilities. Such factors need careful consideration in arriving at a specification threat level. In any case the assessment of the threat to SATCOM should be seen in the context of the threat to NATO communications as a whole: Only if a common philosophy is adopted towards threat evaluation will the relative vulnerability of SATCOM vis-a-vis alternative communication media be properly represented.

5.2 UPLINK JAMMING

This type of attack is attractive to an enemy because many links can be affected simultaneously. Its effectiveness depends primarily on the EIRP achievable by the jammer and thus the most serious form of this threat comes from large static installations.

Of the factors that contribute to the EIRP of a large static jammer, the transmitter power achievable is believed to be limited primarily by technological considerations whereas the antenna size is more litely to be constrained by cost and vulnerability considerations. Having said that, it must be remembered that very large antennas already exist (e.g. for radio-astronomy) whose gains approach the limit of technical feasibility and which could in principle be modified for use as jammers in time of war.

Maximum transmit power levels have been increasing at roughly 10dB per decade and there do not yet appear to be any fundamental technological barriers to further progress. At EHF, the development of very high power gyrotrons for use in nuclear fusion research is of particular concern. However, the problems of connecting such a device to an antenna should not be under-estimated.

Mobile (e.g. shipborne) and transportable jammers will be limited principally by primary power and pointing accuracy. While the achievable EIRP levels are much lower than the maximum threat from large static jammers it is argued by some that the latter, being very conspicuous, would be relatively easy to destroy in war and that mobile and transportable jammers, being much more resilient and numerous, would pose a more significant threat. Of course, many such jammers could operate in unison to raise the effective jamming EIRP by a large factor.

Uplink jamming from airborne platforms poses a relatively insignificant threat to a high-altitude satellite because of fundamental limitations on antenna size and primary power. A spaceborne jammer, on the other hand, could prove very effective even with very low EIRP, provided it could be manoeuvred close to its target.

5.3 DOWNLINK JAMMING

Downlink jamming has traditionally been regarded as a much less serious threat than uplink jamming. One reason is that (expect in the case of a space-borne jammer) a downlink jammer can normally only attack one SGT at a time and must get close enough to it to have a line-of-sight path. Despite this, the jammer can have such an enormous range advantage that it can often be effective with only simple hardware and low power levels.

Airborne platforms can be very effective in downlink jamming because of mobility, facility of deployment and large coverage areas.

Downlink jamming from ships is considered a relatively insignificant threat since the need to be close enough to have a line of-sight path involves serious risk of destruction.

Space-borne jammers in low earth orbit have limited effectiveness because of their short viewing time of the target SGT's main beam. A downlink jammer satellite in geostationary orbit could be effective if manoeuvred close to a NATO COMSAT. However, alternative forms of attack from such a platform (e.g. physical attack, uplink jamming) are considered more realistic.

In one possible scenario, simple low power jammers capable of being carried by a man are scattered in the region around an SGT by clandestine forces. Jamming through wide-angle sidelobes could be effective for several hours while the jammers were located and disabled. However, it can be argued that such clandestine forces would be well placed for a physical attack on the terminal, which would be more permanent in its effect.

5.4 INTERCEPTION

For certain terminals, there may be requirement for covert operation using Low Probability of Intercept (LPI) techniques. Essentially these consist of wide-angle sidelobe suppression and spread-spectrum modulation to lower the spectral density of the transmitted signal. LPI terminals must also, of course, use the minimum transmitter power consistent with the required data rate and the current jamming level. The threat is twofold; firstly that by detecting the presence of a signal an enemy may locate and destroy the terminal, secondly that by analysing the intercepted signal he may deduce something useful about message content, modulation etc. which will yield intelligence data and may also help him to jam or "spoof" the system.

The interception threat is determined by the enemy's strategy in searching the area believed to contain the terminal, the type and range of the interception platform and the number and sensitivity of the receivers employed.Interception by airborne and spaceborne platforms is likely to be most effective because of the wide field of view. The wideband radiometer is likely to remain one of the most effective means of detecting spread-spectrum signals: As the probability of detection of this device depends directly on signal strength but only on the square root of the dwell time, it can easily be shown that it is more effective to search a given area in a given time using a narrow beam than using a wide beam. Thus a major factor in the interception threat will be the enemy's ability to deploy and point high gain antennas from airborne and spaceborne platforms.

As the probability of detection by a radiometer decreases with signal bandwidth the use of very wide bandwidths, e.g. 2 GHz at EHF using frequency-hopping, is attractive for LPI applications.

However, the interceptor can counter this by channelizing his receiver, at the cost of extra complexity. Use of digital techniques together with VLSI will progressively increase the degree of channelization that is feasible. Even without such refiments, the prospects for LPI operation are not good: It is shown in [5.2] that even by exploiting the wide bandwidth and high atmospheric attenuation at EHF, special measures must be taken to screen the terminal in order to prevent detection from airborne platforms. Even then, the introduction of jamming at only a very low level-possibily one at which its true nature would be hard to recognise-would force the terminal to increase its power to a level where detection could readily occur.

The use of "transparent" satellite channels makes it possible for an enemy to monitor individual satellite accesses via the downlink. This may be done from a terminal-possibly transportable anywhere within the footprint, and avoids the cost of spaceborne or airborne interception platforms. The interceptor may also be able to locate the transmitting terminal by precise frequency measurement and knowledge of the satellite's orbit, for example.

These dangers can, of course, be avoided through the use of on-board processing.

5.5 NUCLEAR THREAT

A comprehensive study of the nuclear threat to NATO SATCOM has recently been carried out by SHAPE Technical Centre. The findings are reported in [5.3], and are summarised briefly here.

Nuclear effects are categorised as follows: Effects on the propagation path, effects on the space segment and effects on the ground segment.

5.5.1 Threat to the Propagation Path

The most serious nuclear threat to the propagation path comes from high altitude nuclear detonations, which can cause serious degradation over very large areas of the earth's surface because of the extended effects in space and time. The same class of detonations also cause a strong electro-magnetic pulse (EMP) to the ground segement within line-of-sight of the burst.

The nuclear detonation effects on the propagation path can be divided into "non-scintillation effects" and "scintillation effects".

Non-scintillaton effects result from enhanced total electron content of the propagation path and can disrupt communications via the links interdicted by the fireball. These effects become dominant during the first minute or so from burst and cause degradation in the system performance by the following mechanisms:

- Absorption
- Increase in antenna noise temperature
- Fluctuations in phase and frequency
- Delay jitter
- Dispersion
- Ray bending

In contrast wiht non-scintillation effects, the degrading effects of scintillation last much longer and constitute the major theat to the propagation path. The degration can be outlined in three main areas:

i) Fast and Slow Rayleigh Fading:
The decorrelation time,which is used as a measure of the fading,increases with frequency and with time-after-burst.

ii) Frequency Bandwidth Limitation:
The frequency-selective bandwidth of the channel,also increases with frequency and time-after-burst. The frequency dependence is much more pronounced that that of (i).

iii) Angular Scattering Loss:
The loss suffered by a high-gain antenna as a result of angular scattering is related to the beamwidth and therefore increases both with frequency and antenna size. The loss decreases with time after the burst.

In the early phase of scintillation, fast fade failure occurs because of the rapid phase changes in the received signals. Consequently, the demodulator cannot distinguish between phase changes due to modulation and scintillation. The maximum degradation is anticipated when the decorrelation time becomes comparable with the data rate. It is expected that the lower data rates (e.g. telegraph) will be the most strongly affected by this mechanism.

Later,with increasing decorrelation times, amplitude scintillation, bounded by slow Rayleigh fading, becomes the dominant source of degradation. During this period, which lasts longer, one observes burst errors during fades, acquisition and tracking difficulties and a decrease in channel capacity. The degradation depends on the system and the employed modulation scheme.

Mitigation techniques include coding with long interleaving depth, improving synchronisation performance, increasing the link margin and using diversity.

The limitation on the maximum bandwidth usable by SATCOM links is a result of frequency-selective fading, i.e. scattering of multi-path components with randomly varying delays. Systems using instantaneous bandwidths larger than the one permitted by the disturbed path undergo serious degradation. In this respect, direct-sequence spread-spectrum systems are more vulnerable than frequency-hopping systems. Inter-symbol interference constitutes the maim source of degradation, with a resulting irreducible error rate. Mitigation is almost impossible by increasing the EIRP. A reduction in the chip rate can decrease the inter-symbol interference, correlation loss and demodulation loss but with a resulting decrease in AJ performance Use of large receiving antennas also helps, though with an accompanying increase in angular scattering loss (discussed below).

A high-gain receiving antenna attenuates, with its narrow beam, the signal components arriving from off-boresight directions. Thus the degrading effects of signal decorrelation and bandwidth limitation can be partly alleviated. However, because of the rejection of these components the received signal energy decreases. This decrease is accounted for by the angular scattering loss of the receiving antenna. The angular scattering loss is significantly large and long-lasting at SHF. It increases with the square of the antenna size and the fourth power of frequency. These effects can be mitigated against by increasing the EIRP and by smaller-sized receiving antennas.

5.5.2 The Threat to the Space Segment

The nuclear radiation effects on satellites can be divided into two categories.

i) Prompt radiation effects
ii) Trapped radiation effects

The prompt radiation lasts less than one second from burst and affects the satellites in line-of-sight by x-rays, gamma rays and neutrons. The main degradation to the satellite comes from the very high dose rates. Prompt radiation can also interfere with the operation of optical detectors, causes darkening in optics, degrades the sensor performance and decreases the signal output. The effectiveness of prompt radiation increases with decreasing separation distance between the satellite and the burst. Protection against collateral attacts, which can target multiple satellites, is very difficult and expensive.

Trapped radiation can affect satellites beyond line-of-sight,

and its effects last much longer than those of prompt radiation. The main degradation caused to the satellite is by high accumulated doses, which lead to rapid "ageing" of the victim satellite.

5.5.3 The Threat to the Ground Segment

The effects of nuclear detonations on the ground segment change fundamentally with the burst altitude. Low altitude detonations can be very effective against only a single SGT because of the localised nature of the effects. However, the effects of high altitude detonations cover areas within line-of sight of the burst. For example, a 150 km burst height is enough to cover the whole of western Europe. The main effect of high altitude detonations is a strong EMP, which can cause serious degradation by inducing very large currents and voltages on conductive material on the earth's surface. Serious degradation is anticipated in unhardened SGTs and associated facilties.Besides, these detonations also disturb the propagation path. Hardening the ground segment against EMP effects should be considered in the context of balanced survivability of the NATO SATCOM system against the spectrum of threats.

5.5.4 Conclusion

In summary, high altitude nuclear detonations can disturb the propagation path and cause a strong EMP to the ground segment within line-of-sight of the burst. This is evidently a very serious threat to NATO SATCOM. Nuclear attacks on satellites are more expensive and some countermeasures may be taken to reduce the weapon's effectiveness. Saturation of the geostationary; nuclear radiation environment by a suitable choice of weapon and burst location may be very effective in ending the useful life of a satellite by rapid ageing.

5.6 PHYSICAL ATTACK

The types of threats which may exist and which system designers may need to consider in the time frame of interest can be classifiela as follows:

Nuclear Burst Effects

ASATS
- Direct Ascent
- Co-orbital
- Kinetic
- Fragmentation
- Nuclear

Directed Energy Beams
- Laser
- Neutral Particle
- High Power RF
- Ground Based
- Space Based

ECM (Jamming)
- Ground Based
- Space Based
- Uplink Jamming
- Dowling Jamming

5.6.1 Nuclear Burst Effects

Satellites may have to contend with direct nuclear attack on the satellites themselves or with collateral effects of nuclear bursts in the outer atmosphere due to, for exampla, ABM weapons or deliberate EMP producing weapons. The damage producing phenomena are the same in both cases; the difference is in their relative intensities.

The primary damage producing phenomena are: X-Rays. Gamma rays, Neutrons, radioactive debris and EMP. X-Rays in sufficient quantities produce surface damage attendant with mechanical shock. They also produce SGEMP (Self Generated EMP), which in turn can cause burn-out or degradatioan in solid state electronics devices. Gamma rays and neutrons penetrate the spacecraft giving rise to so-called TREE (Transient Radition Effects in Electronics) effects which can cause latch-up and contribute to the damage to solid state devices. Radioactive debris refers to the enormous increase in radioactive particles in space over natural levels. Even if the satellite survives the immediate blast, these debris accelerate the degradation effects of the natural environment and limit lifetime accordingly. EMP arises from the impingement of ionizing radition from the blast on the molecules of the upper atmosphere. This pulse, which is usually dispersed by traversing the intervening plasma, can couple high voltages into the satellite electronic circuits. Self Generated EMP, on the other hand, arises in the satellite itself due to the impact of the prompt X-Rays.

All of these nuclear effects are relatively well known and understood. There are design techniques and processes, collectively known as "hardening", which ameliorate these effects and can be applied to a greater or lesser degree depending on the level of hardness desired and the funds available. These techniques include selection of electronics components, selection of materials, radiation shielding, electromagnetic shielding and current limiting measures in the circuit designs. Implementing and validating these techniques will have a noticeable effect on the satellite mass. Perhaps of more concern is the cost impact which is significant.

In the time frame of interest, the situation regarding nuclear effects and countermeasures is not expected to change drastically as it is set primarily by the laws of physics. The advent of new components such as GaAs devices, which tend to be more radiation tolerant than silicon devices, may improve the situation somewhat, but probably not to first order.

5.6.2 Directed Energy Beams

These threats can be classified as to the type of beam laser, particle or RF and as to whether the weapon is ground based or space based.

5.6.3 Ground Based Laser Beam Weapons

In the time frame of interest, it may be possible to generate, steer and propagate through the atmosphere, high energy optical beams that would pose a threat to satellites. A laser weapon has characteristics unlike those of an ASAT. It has an unlimited number of rounds to fire and its attack is virtually instantaneous. On the other hand, it is expected that these weapons would comprise very large, fixed installations, which would be vulnerable to attack themselves. Also, they would not be capable of penetrating clouds, which would limit their flexibility in a battle scenario other than a surprise attack at a time of choice.

Propagation through the atmosphere requires the use of atmospheric compensation technology as was demonstrated on the Space Shuttle flight several years ago. Even with such compensation, it is not expected that the energy could be focused to a small spot on the satellite, but rather would irradiate large portion of the satellite or its entire surface with its beam. Satellites in low earth orbit could be illuminated with exceedingly high flux densities lasting from seconds to minutes. At these high power levels, the damage mechanism would be simply thermal load, which could damage or even vaporize surface materials thereby causing destruction to the satellite. Protection by reflective coatings would not be effective as only a small absorption of energy would be sufficient to overheat the material. It might be possible to shield the satellite at great expense in weight by a highly insulating, refractory material such as "shuttle tile" which would absorb and reradiate the energy at a very high surface temperature while insulating the satellite from this

temperature for a sufficiently long time to get out of range of the weapon.

In higher orbits the threat is considerably less severe. The energy flux density decreases as te square of the range, and it is unlikely that laser weapons will be a cost effective means for destroying high orbit s atellites in the time frame of interest. However, lasers could pose a threat short of destruction by attacking and upsetting optical sensors or by causing transients in solar panels. These are factors that the satellite designer may need to consider, but on balance it is not expected that laser weapons will be a fundamental threat to satellites at near-synchronous altitudes in the time frame of interest.

5.6.4 High Power RF Beams

Of course RF beams pose a jamming threat, but it may also be possible to cause permanent damage to receivers or other sensitive electronics by very high power microwave sources such as the gyrotron, which has been under development for some years, particularly in the Soviet Union. Like the laser, the energy density decreases rapidly with range. For near-synchronous satellites, therefore, it is expected that, if such a threat did develop in the future, one could counter it effectively by design techniques such as filtering and blanking the receiver to protect it against the incident in-band flux. This would render the RF weapon no more effective than a straightforward jammer, therefore, it is concluded that the jammer is the much more plausible threat on cost-effectiveness grounds.

5.6.5 Particle Beams

A particle beam weapon would have many similar characteristics to a laser weapon in terms of its size, the damage mechanisms and the fact that increasing altitude works to the satellite's advantage. The beam would have to consist of neutral particles such as atoms or neutrons because the earth's magnetic field will deflect charged particles excessively. Even though such beam weapons have been considered for the American SDI program, it appears that the technical problems are even more severe than those associated with laser weapons. It is concluded, therefore, that this weapon will not be a dominant threat in the time frame of interest unless an unforeseen technology breakthrough occurs in the meantime.

5.6.6 Space Based Directed Energy Weapons

It was mentioned above that the inverse square law decreases the effectiveness of beam weapons for high orbit satellites. This range advantage could be circumvented by the attacker if he bases his weapon in high orbit, also so that the power levels needed for damage or destruction could be much lower than for ground based weapons. Space based beam weapons were considered in some depth for the SDI program. There it was concluded that kinetic weapons were superior given today's technology and reasonable extrapolations thereof. Mainly on this basis, WG-13 also concludes that space based beam weapons are unlikely to be a significant factor in the threat equation for the frame of interest. However, space based threats of other kinds such as jammers of ASATS could indeed pose a significant class of threat.

5.6.7 Anti-Satellite (ASAT) Weapons

ASATs can be classified as to their method of closure with the target satellite-direct ascent or co-orbital-and as to the type of warhead-kinetic kill, fragmentation of nuclear. A direct ascent ASAT can be launched from the ground, the air or from a low orbit. It is fired on a ballistic trajectory that intercepts the target satellite, usually at extremely high velocity. The use of a homing sensor and terminal guidance is assumed. A co-orbiting ASAT attacks in a more leisurely-but perhaps more covert-manner by attaining the same orbit as the target and essentially rendezvousing with it and detonating its warhead.

The kinetic kill warhead destroys the target by simply colliding with it; at typical impact velocities the release of energy is tremendous. This type device is applicable to the direct ascent attack scenario. The fragmentation warhead, on the other hand, is a high-explosive device set off in proximity to the target and killing it with shrapnel. It may be used in both direct ascent and co-orbital attacks. The nuclear warhead is similar, but of course very much more lethal and is effective to ranges of many kilometers.

5.6.8 Direct Ascent ASAT's

The direct ascent ASAT is technically quite feasible, certainly in the subject time frame. Missile propulsion systems and guidance systems exist today which probably could do the job against satellites in any orbit if sufficient financial resources were to be devoted to such ends. Against a direct ascent, kinetic kill ASAT, shielding is ineffective due to the high impact velocity and the mass involved. Defencive tactics must, therefore, include such measures as maneuver, decoys, flares and similar methods that exploit weaknesses in the guidance and control systems of the ASAT. Again, the increased range of high orbit works to the satellite's advantage because even a direct ascent takes several hours which gives time to take evasive actions if a warning of the attack is available. A necessary ingredient of effective defense against this class of ASATs is therefore a space surveillance and warning system that will alert the satellite of impending attack. Given the ever increasing emphasis on the military use of space. it is reasonable to expect that some form of national space surveillance and warning system will exist in the time frame of interest and that NATO will have access to such warnings.

Another effective tactic against this threat would be proliferation. A direct ascent ASAT attack will obviously be an expensive undertaking. If the value of any individual target is relatively low and the number of targets is relatively high, then this kind of ASAT attack is not profitable to the enemy.

5.6.9 Co-Orbital ASAT's

The working group considers this to be a feasible and likely threat in the future. The nature of this attack is quite different from the direct ascent case. Here range is of no advantage to the target satellite, especially if the attack is covert. The ASAT could lie undetected near its quarry and explode on a moments notice. This is sometimes called a space mine. Again, however, attack warning can make defense feasible and is essential to the defense. If it is known that a co-orbital ASAT attack is underway, a number of defensive options are available. These include evasive maneuver, shoot-back and retaliation against the antagonist's communications satellites. Proliferation is also expected to be effective, although not so effective as in countering direct ascent attacks.

5.7 PIRACY (Unauthorised Access)

A transparent transponder may be accessed with communication signals by unauthorized users (pirates). If the pirates are cabaple of using spread spectrum accesses, these can take place without being detected by NATO SATCOM operating authorities.

The future satellites will depend on on-board processing (see Section 8.2) and use transparent transponder only to provide overlap capability with earlier phases of the ground segment.

To access a processing transponder of any kind, including the future high-capacity processors according to the principles of Section 8.2 will depend on knowledge of the TRANSEC code. Anyone able to use the satellite as a pirate is therefore also able to deny the authorized users the jamming protection of the processor. The concequences would of course be catastrophic and this situation must be avoided.

Piracy exploitation of the satellite does not represent a problem

in addition to protection of the TRANCES code and other measures used to obtain and maintain AJ properties.

5.8 REFERENCES

[5.1] Şafak, M., "An Assesment of the Jamming and Interception Threats to the NATO SATCOM System", STC Technical Memorandum TM-872 (Draft) (NATO Secret).

[5.2] Masterman, P.H., "Initial Considerations for a Post-2000 NATO SATCOM System", STC Technical Memorandum TM-886 (Draft) (NATO Secret).

[5.3] Şafak, M., "An Assesment of the Nuclear Threat to the NATO SATCOM System", STC Technical Memorandum TM-884 (Draft) (NATO Secret).

CHAPTER 6

THREAT IMPLICATIONS AND COUNTER-MEASURES

6.1 ECCM TECHNIQUES

6.1.1 The Concept of Survivability as Applied to ECM/ECCM

There are several factors which can be optimised for low cost and high performance but this must be done in the full system context and by paying simultaneous attention to all aspects of the system including the environment in which the system operates, ie: the jamming environment. Since multiple interactions occur in the establishment of a design, iterations are employed. The implicit structure for antijam design iterations is shown in Fig.6.1. The point at which iteration is stopped depends on the criteria used for performance anlysis, ie: definition of survivability against jamming. For a defined threat, survivability is defined as achieving either satisfactory or mimimum communications in the face of the given jamming level. If, however, we assume no knowledge of either actual or postulated threats on the premise that any communications system can elicit a technological counter response which, if succesful, requires redesigning of the system which, in turn elicits another counter response. This escalation leads to three different defitions of survivability:

a) Concept of cost exchange ratio defined as
<u>Enemy cost to negate</u>
own cost to deploy

If the cost ratio is greater than one, the system, or an incremental improvement, may be said to be survivable.

b) Another criterion relates to the operational viability of the postulated counter strategy. If the system is designed to demand a counter response that is either too visible, too difficult, or of uncertain success, it may not be adopted. This attribute of forcing responses judged unreasonable, is the second criterion of survivability. Unreasonability of required threat can be:

- High visibility
- Impossibility to move
- Nonsanctuary location demanded
- High number of jammers
- High absolute cost

If (a) and (b) are both satisfied with respect to jamming, the enemy may choose to physically attack the satellites. This leads to:

c) The principle of "balanced survivability":

Cost of given attack is approximately equal to cost of any attack (jamming).

The most desirable designs are those satisfying all three criteria (a), (b) and (c).

6.1.2 Factors Contributing to AJ Performance

In general, a jam resistant system employs some form of bandspreading signalling and processing (demodulating repeaters) and/or null-forming spacecraft antennas. Under these circumstances, Jammer-to-Signal ratio (J/S), also called processing gain, is given by:

$$J/S = (\psi)\left(\frac{W}{R}\right)\left(\frac{1}{E/N_o}\right)$$

where

ψ = Space Spreading Factor
W/R = Band Spreading Factor
$1/E/N_o$ = Modulation and Coding Factor

There are very large number of possible distinct systems designs which must be quantitatively defined and studied before arriving at the optimum architecture for future defence satellite communications systems. Table 6.1 below gives some limits for the three factors mentioned above and shows that the band-spreading is an important area where high gains against jamming can be made. In this respect the use of frequency-hopping spread spectrum technique holds particular promise for very wide spreads.

Table 6.1

SURVIVABILITY LIMITS

Factor	Lower Bound	Upper Bound
ψ	0 dB	10-30 dB
W/R	43 dB	100 dB
$1/E/N_o$	- 15 dB	-5 dB

Spread spectrum techniques are discussed in Section 6.1.3.

The second main AJ technique is spatial discrimination by means of antenna techniques. In milsatcom, the major use of spatial discrimination is for spacecraft receive antennas for the uplink. Simple spatial discrimination includes the use of low sidelobe spot beams so that jammers outside the spot area are attenuated and the use of fixed nulls over known hostile areas such as over WP countries. More complex spatial discrimination involves adaptive antenna nulling and is discussed in Section 8.3. Spatial discrimination can also be used by terminal downlink receive antennas. Generally, a narrow beam with low sidelobes gives sufficient spatial discrimination. In extreme conditions, adcaptive nulling may be needed. However, at SHF and above, such a nulling capability need only be simple since it would be unlikely to have more than one or two jammers deployed against a particular terminal. Therefore, simple sidelobe cancellers would likely suffice.

As mentioned earlier in this section, the two AJ methods can be combined to obtain the benefits of both. If the processing again of the spread spectrum is PGss and the null depth of the antenna is ND, then the combined benefit is approximately $ND.PG_{SS}$ or, in dB, is $10LogND + 10LogPG_{SS}$. In systems under current development, values of PGss of 50 dB and higher and values of ND of 30 dB and above are feasible. The trend appears to be to start with SS alone as the principal AJ method since it is usually less costly, is more technologically advanced, and can provide more AJ protection. However, if the jamming threat level exceeds the capability of the SS alone, then nulling is added.

6.1.3 SS Techniques for AJ

<u>6.1.3.1 Introduction</u>

Spread Spectrum (SS) techniques for purposes of AJ are well known and have been described extensively [6.1]-[6.3]. SS technology is well developed. It has been utilised on numerous existing systems such as NATO3, SKYNET3 and 4, LES8 and 9,

FEP, Syracuse I, etc. It appears that SS will be a standard feature for milsatcom systems. In SS, the signal to be transmitted is intentionally spread over a frequency band that is much wider than the information bandwidth. At the receiver, this wideband signal is compressed back to its original information bandwidth. The idea behind this process is to force a jammer to spread its energy over this very wideband, thereby diffusing its effect. A potential processing gain is sometimes defined and is given by

$$PG_{ss} = W_{ss} / R_b \quad (1)$$

where R_b is the information bit rate and W_{ss} is the spread bandwidth. This PG is only valid against a jammer using white gaussian noise across the entire spread band. Unfortunately, simple jammer strategies (EC3M) that the communicator can make to overcome these jamming strategies that either gives a PG_{ss} equal to (1) or, equivalently, forces the jammer back to using wideband noise.

There are three basic types of SS waveform: direct-sequence (DS), frequency hopping (FH) and time hopping (TH). There also hybrids of these with DS/FH being the most common. One way to explain the difference between the three basic types is through a time-frequency domain of Fig 6.2 in such a form and manner as to cause the most degradation in the communications error rate performance.

6.1.3.2 Direct Sequence

a) Introduction

A simplified block diagram illustrating implementation of DS SS is shown in Fig. 6.3. Data at bit rate R_b is added modulo-two to a pseudo-random (PN) sequence. The bits of this sequence, often called chips, have a chip rate, R_c, that is much higher than R_b, thereby resulting in the spreading of the spectrum. The resulting sequence is usually modulated with some form of PSK and the resulting spectrum illustrated in Fig. 6.2. is continuous in time with a spread bandwidth of W_{ss} which, at the 3-dB points is about R_c.

At the receiver, an idential replica of the PN sequence is multiplied with the incoming signal in such a way that the output is the desired data sequence at the original data rate, R_b. In terms of frequency, the spread bandwidth, W_{ss}, is compressed back to the bandwidth of the data. Any interference also goes through the multiplication process so that its spectrum is its received spectrum convolved with the spectrum of the PN sequence of width W_{ss}. This process spreads the interference over a band $>W_{ss}$. A narrow band filter,with a width of the order of R_b,passes all the data but rejects most of the interference thereby providing an AJ capability.The processing gain is against wideband noise,then,is given approximately by(1).

b) DS- Transec

In the simplest form of DS SS,the PN chip sequence is the same for every data bit transmitted.Such operation greatly simplifies the implementation. Unfortunately,such a repeated sequence is easily detected by an enemy who can then use it in a jammer to totally negate the advantage of using DS SS.Therefore, it is more common to use very long PN sequences so that a chip sequence within a data bit period never repeats. Synchronization for such sequences is,of course,quite a problem and much effort goes into solving it [6.1,vol.3].

If the PN sequence is generated by linear feedback shift register containing N delay elements,then interception of 2N contiguus chips enables an enemy to calculate the feedback taps and current state of such a sequence generator. Therefore, the enemy is able to calculate all future chips,thereby permitting the enemy to jam with the correct sequence and negating the AJ capability.Thus,some form of protection,sometimes called "Transec",against breaking of the sequence is needed. One way is to encrypt the sequence from the linear feedback shift register. Furthermore, the key for such encryption will normally be changed periodically such as on a daily basis.

c) DS-EC3M/EC4M

The easiest EC^3M strategy of a jammer to reduce the effectiveness of the AJ capability of DS SS is to use a CW tone at the center frequency of the DS spread spectrum [Sec 2.5 of 6.2]. This form of jamming is somewhat more effective than wideband jamming [Sec. 2.4.6 of 6.3] and is probably the easiest jamming signal to generate at high power levels. There is considerable literature on the topic of combatting tone jamming by means such as adaptive notch filtering. Unfortunately,such tone nulling techniques can in turn be countered by using multiple tones,and tones with jittered center frequency.

If the jammer discovers that the communicator has employed anti-tone techniques, then there is an alternative EC^3M strategy. This alternative is for the jammer to generate its own DS spread signal at a carrier and chip rate as close as possible to that of the communicator's DS [Sec 2.5 of 6.2]. This jammer's sequence does not match that of the communicator.Nonetheless,this strategy degrades the performance of the DS SS system somewhat more than the same level of wideband noise jamming and there appears to be no effective counter to it.

d) DS Advantages/Disadvantages

DS SS has found good use where the spread bandwidth, W_{ss} is not excessive with chip rates, R_c, of 1 to 10 M chips/s being typical. Values of R_c beyond 100 M chip/s greatly strain the technology required for generating PN sequences and make synchronization extremely diffucult. Furthermore, above 100 M chip/s dispersion in the SATCOM propagation path can degrade performance.

DS SS can also be employed for multiple access purposes in addition to its AJ purpose. In the multiple access application each user pair is given a different PN sequence "code" whence the name "code division multiple access" (CDMA). Regrettably some inappropriately have used the term CDMA to mean DS SS even if no multiple access is involved. This incorrect and confusing terminology has been carried over to a certain extent to NATO.

Since the processing gain is limited by the maximum R_c, and there are jammer strategies that further reduce the effective processing gain, it is concluded that DS is best employed in AJ applications where jammers are relatively unsophisticated. In general, DS is not appropriate to the Soviet jamming threat.

6.1.3.3 Frequency Hopping

a) Introduction

A simplified block diagram illustrating implementation of a fast FH SS transmitter is shown in Fig 6.4 and the corresponding receiver is Fig. 6.5 [6.4]. The definition of "fast" used here is relative to the hop rate and is defined as fast if there are one or more hops per transmitted information symbol. Initially, only fast FH is considered.

For milsatcom, fast FH generally uses non-coherent M-ary FSK modulation. Such M-ary symbols are created by generating one of M possible tones at a symbol rate of R_s. This symbol tone is then mixed with the hopping carrier frequency, hopping at a rate $R_h > R_s$. The hop frequencies are generated from a pseudo-random sequence so that they are uniformly distributed across the spread band W_{ss}. A typical frequency-time pattern is shown in Fig. 6.2 where the transmitted M-ary tone has a duration T_h and an instantaneous bandtwidth of $1/T_h$. Usually $R_h = 1/T_h$. We define an M-ary channel as consisting of M

frequency bins. Each bin has a width R_h.

At the receiver, Fig. 6.5, the received signal is dehopped by means of a hopping LO that has the same hopping pattern as the received signal. The dehopped signal then goes to a demodulator which has the form of a bank of matched filters. These can be implemented in a number of ways as discussed in Sec. 6.1.2

The processing gain against a wideband noise jammer is not given by (1) as it is sometimes erroneously thought. rather it is

$$PG_{fh} = W_{ss} / R_h. \quad (2)$$

Since $R_h > R_s$, then PG_{fh} is not as large as predicted by (1). It will be shown later how this deficit can be mostly recovered.

b) FH-Transec

If the hopping frequencies are defined by the PN sequence in a linear, sequential, and unscrambled order, and furthermore if the PN sequence is generated by a linear process, then it is possible for an interceptor to measure the hop frequencies and determine the PN generator's present and future states. Therefore, it is then possible to perform jamming only at the hop frequencies thereby eliminating the AJ capability of the FH SS.

To overcome this threat, some form of nonlinearity and scrambling is inserted between the linear PN generator and synhesizer. Some very simple but robuts processes can be used for this Trances function. However the security organizations in some member countries have mandated that this function be implemented by an official encryptor. Fortunately, for this application, only encryptors are needed at either end, i.e. no decryptors are needed, greatly simplifying the operation.

c) FH-EC3M/EC4M

The processing gain (2) can be greatly reduced by a simple EC3M jamming strategy. Instead of spreading the total jamming power, J, over the entire hopping band at a power density of J/W_{ss}, the jamming is restricted to a fraction, γ, of the band but at a power density increased by $1/\gamma$. On the fraction $1-\gamma$ of the hops, there is no jamming. On the jammed hops, the error rate can be so high that even averaged over the unjammed hops, the overall error rate is unacceptable. For example,if 1/10 of the hops have a bit-error rate of 0.5, the worst possible, the overall bit error rate is 0.05 which is very poor. The jammer attempts to choose the value of γ that gives the worst-case value of overall error rate [6.1], [6.4]. It is a trade off between maximizing the error rate in the jammed hops and maximizing the number of hops jammed.

The form of the jamming is often considered to have two principal forms. One is partial-band noise (PBN) jamming wherein gaussian noise is used and spread over the band γW_{ss}. The other is multiple-tone (MT) jamming wherein discrete frequency tones are spread over some part of the band. In principle, MT jamming can cause somewhat more degradation in performance than PBN jamming [6.4]. In particular, in what is called "worst-case jamming in Houston's sense" [6.4], the spacing of the tones are chosen so that only one jamming tone would ever occur within an M-ary channel on any hop and that this tone slightly exceeds the signal in amplitude. Such jamming can be seen to optimize the jammer's resources by putting precisely just enough energy into one channel to cause a symbol error. In order to compare FH systems of differing hop band widths,it is usual to define an effective signal-to- jammer ratio by [6.4]

$$SJR = \frac{E_h}{J_o} \quad (3)$$

where E_h is the received eenergy per hop and J_o for noise jamming is the power spectral density that would arise if the power were uniformly spread across the hopping band. Although a CW tone does not have a spectral density defined, it is nonetheless common to define an equivalent uniformly spread J_o.

Some of the above issues can be understood with the aid of Figs. 6.6. and 7 where the bit-error probabilty is plotted as a function γ of bins jammed or for Fig. 6.7 the fraction, β, of channels jammed. Three values of SJR are used. Fig 6.6 is for PBN jamming and Fig. 6.7 is for MT jamming with only one jamming tone per channel. In both Figures curves with and without system noise at a level of SNR=13.35 dB are given. The break point for MT jamming when no system noise is present is the point at which the jamming tone exactly equals the signal tone; for β exceeding this value, no errors can occur.

It is seen that for the same SJR, the MT jamming causes the highest error probability. For both PBN and MT jamming, there is always some value of γ that gives a maximum error rate which is denoted "worst-case" jamming. For large SJR,the region near the worst case is a narrow peak which broadens as SJR decreases.For low SJR,the peak is at or near $\gamma=1$ in Fig.6.6 and $\beta=1$ in Fig. 6.7.

At SJR>20 dB,the worst case jamming peak results in a bit-error probality of about ≤0.01.Such error rates can be handled by error-correction(EC) coding. Coders and decoders are shown in Figs. 6.4 and 6.5, respectively. Under these SJR conditions, the EC coding has the appearance of providing enormous processing gain. For example, a modest EC code can correct an error rate of 10^{-2} down to 10^{-5} whereas to achieve 10^{-5} by means of increasing SJR would require, under worst case jamming, and increase from 20 dB to 45 dB with an apparent processing gain at an astounding 25 dB! Such gains are problematic since it assumes a jammer could actually determine the exact worst-case value of γ or β at high SJR where the peak is very sharp.

For bit-error rates above about 0.01, EC coding becomes less useful. In fact, some EC codes such as the popular rate-1/2 constraint length-7 convolutional code, the decoding can actually fail and give an error rate out of the decoder that is higher than the input error rate. Fortunately, a method usually called "diversity combining" can take input error rates that are $>10^{-2}$ and correct it to have error rates $<10^{-2}$ at the expense of data rate. Often, as shown in Fig. 6.5, a diversity combiner is concatenated with an EC decoder so that the diversty can bring the raw error rate down to level that the EC decoder works effectively.

For the diversity methods, the transmitted symbol is repeated on L hops thereby providing a redundance of L but with a symbol-rate reduction by 1/L. The simplest form of diversty combining is hard-decision majority-vote (HDMV) combining [6.4.]. In HDMV combining, a hard decision as to which of the M possible symbols was received is made on each of L hops. The frequency bin with the most count is declared to be the symbol received. This technique is simple to implement, can have an easily charged diversity level, L, and performs reasonably well against PBN jamming or MT jamming at moderate levels [6.4],[6.5].

A diversity combining method must not only correct well at low SJR but also be robust to changes in jamming strategy. Numerous diversity combining techniques correct well under certain conditions but fall apart and perform very poorly under changes in jammer type. Since the communicator has no control of the jamming strategy, such combining techniques are to be avoided.

One diversity combining method that performs well at low SJR and is robust to jammer strategy is normalized envelope detection (NED). NED appears to have been described first by Gong of Raytheon in the U.S. Since then, a number of similar

methods have been proposed that perform about as well as NED. In NED combining, on each hop the amplitude of the envelope of each of the M frequency bins is divided by the sum of all M envelope amplitudes. Then the sum over L hops of each normalized bin is taken and the largest chosen as the symbol received. This method is particularly useful against large tone jammers. In an experimental system [6.5], it was found that NED can be implemented in real time with a DSP chip and it indeed can correct high input error rates for both PBN and MT jamming. Furthermore, the result in [6.5] showed that the discrepancy between the ideal processing gain of (1) and the actually FH processing gain of (2), could be largely gained back by the NED combining. This discrepancy, $L=R_h/R_s$, can be viewed as the coherent combining loss i.e. it is gain that would accrue if the hop duration were expanded from T_h to LT_h. Thus, we see that any other diversity combining method is not likely to be much better than NED since it already approaches the coherent combining gain.

It is cautioned that the hop rate should not be reduced by 1/L in an attempt to achieve the potential 10 Log L dB coherent processing gain. The effective SJR defined by (3) is increased by this cohorent gain. However, cohorent gain is achieved only against noise jamming across the entire hopping band($\gamma=1$). As was seen in the discussion of Figs. 6.6 and 6.7, the jammer can change its γ or type of jamming to negate any benefit of the cohorent gain. In fact, there are simple jamming strategies that can degrade the error-rate performance to worse than if the hop period had not been extended. It can be shown that diversity combining is a far more effective means of utilizing the redundancy L than reducing the hop rate.

Interleaving of data bits or symbols can be used to reduce the chance of bursts of errors which some EC coding methods do not handle well. However, for fast hopping, and PBN or MT jamming, the random hopping ensures that the symbol errors are randomly distributed so that the symbol errors are randomly distributed so that interleaving is not needed. Interleaving, however, does find use for mitigating the effects of fast fading such as caused by nuclear scintilation.

On a philosophical note, the EC4M strategies considered above all have one thing in common. They result in a final error rate that is the same or better than if the jammer had used wideband noise jamming. Thus, the benefit to the jammer of altering the jammer strategy is lost.

d)FH-Slow Hopping

If slow hoppings is used then there are more than one symbol within a hop period. The instantaneous bandwidth is no longer the $1/T_h$ shown in Fig. 6.2 for fast FH. The data modulation for slow FH is not limited to FSK and is often differential PSK. Differential tends to be used because the loss of phase coherence between hops makes coherent PSK less practical. The instantaneous bandwith becomes that of the PSK modulation which is typically approximately the symbol rate, R_s. Thus, the processing gain (2) is no longer applicable and becomes

$$PG_{fh} = W_{ss} / R_s. \qquad (4)$$

Since for slow hopping $R_s >> R_h$, there is much less processing gain available for slow hopping.

Slow hopping is very prone to further performance reduction by the jammer going to PBN or MT jamming strategies. On the hops that are strongly jammed, the bit-error rate is very high resulting in long bursts of errors. Therefore, interleaving is essential to distribute bit errors uniformly so that EC coding can work at its best. Alternatively, codes that are good at detecting error burst and then doing erasures should work equally well or even better because the low quality hops are identified and rejected.

Clearly, slow FH has much inferior AJ protection to that of fast FH for given transmit power. However, slow FH can be very useful in applications where the data rate is higher but the jamming levels are less. Such conditions arise, for example, on the downlink where the jamming threat is less, or on the uplink where antenna nulling is used to reduce the jamming level seen by the receiver.

e)FH-Advantages/Disadvantages.

Remarks here will apply primarily to fast FH. The big advantage of FH over DS is that much wider spread bandwidths, W_{ss}, can be practically implemented. For exmple, 100 MHz bandwidths have been implemented at SHF and 2000 MHz at EHF. Therefore, the potential processing gain, which is proportional to W_{ss}, is much higher for FH than for DS. Furthermore, FH is less prone to further degradation by jammer strategy changes (EC3M) attacks. It is concluded that FH is, overall, much superior to DS for AJ purposes when the jamming levels are large.

Another advantage of FH arises in SATCOM when multiple user signals are to be supported and onboard processing is used. FH is ideally suited to frequency-division multiple access (FDMA) on the uplink and even time-division multiple access (TDMA) is handled. Conversely, DS using CDMA would require separate code generator, demodulator, synchronizer, etc. for each user signal so that it quickly becomes impractical as the number of users increase.

The disadvantages of FH appear to be primarily with the expense in terms of weight, size, and power of the hopping synthesizers. To a lesser extent, there are problems with synchronization but seem to have been succesfully solved in current systems.

6.1.3.4 Time Hopping

In time hopping (TH), time is divided into segments of duration, Ts. The signal is transmitted in a short burst of duration Tp, at pseudo randomly determined location within the segment as illustrated in Fig 6.2. Because of the burst nature, the spectrum width, W_{ss}, is much wider than the information bandwidth R_b.

The use of TH for AJ purposes does not appear to be common. It is surmised that its lack of popularity arises from the fact it is necessary to have a transmitter with a high peak power in order to support the bursts. It is also likely that its AJ performance is limited because the maximum value of W_{ss} is limited by practical constraints such as difficulties in synchronization if W_{ss} exceeds, say 10 MHz.

6.1.3.4 FH/DS Hybrid

Individual hops of an FH system can be spread in turn by DS spreading. Since the potential processing gain defined by (1) is determined by the total spread bandwidth, W_{ss}, there is no additional AJ advantage by the addition of DS if W_{ss} remains equal. Worse yet, there are EC3M strategies that become available to the jammer that degrade performance further that do not appear to have easy EC4M responses as does pure fast FH. Also, the pratical aspects of synchronization for a FH/DS hybrid are certainly much more diffucult than for pure FH. In short, there appears to be no performance advantage and considerable practical disadvantages to using FH/DS hybrids and it is not recommended.

6.1.4 On-Board Processing versus Transponding

Most milsatcom systems to date that use SS AJ have tended to use "bent pipe"transponders on-board-the satellite. The AJ processing gain of the SS can indeed be achieved with such a transponder. However, it has a serious vulnerability sometimes known as power capture which is now explained.

Let the uplink signal power be S and jammer be J. Let the total saturated transmit power of the satellite be $P_T = S+J$ where the gain through the transponder is set at unity. Thus, the signal

power is suppressed by the jammer to be $S=P_T-J$. At the terminal receiver, the ratio of received signal power to system noise, N, is proportional to

$$S/N = (P_T - J) / N. \quad (5)$$

with no jamming $S/N = P_T/N$; it is this ratio that a terminal receiver design is usually based. However, with jamming, the signal is suppressed by the jamming relative to the system noise.

The severity of the problem can be illustrated by a simple example. Suppose that J/S is 10 dB which is a very modest jammer and can be easily handled by SS of modest AJ processing gain. Unfortunately, the power robbing causes the SNR at the receiver to decrease by 10 dB which must be accounted for by adding 10 dB of margin into receiver, which is a severe penalty.

In the face of an AJ SS system on a transponding satellite, a jammer would not bother trying to defeat the SS AJ. Instead, it would be much easier to suppress the signal into its own system noise. Furthermore, the jammer does not need any special waveforms, any will do. For example, a CW tone, which is very easy to generate at high power, will suffice.

It is sometimes argued, that the satellite power amplifier can be operated in a linear mode to avoid the signal suppression discussed above. unfortunately, all that approach does is switch the margin problem from the terminal receiver to the transponder HPA.

Onboard processing (OBP) that does at least demodulation and remodulation eliminates the power robbing problem. In the face of strong jammers, the use of OBP appears essential. OBP will become a common feature on future milsatcom systems.Appendix 6A contains an analysis method for ECCM links via both processing and hard limiting transponders. Applications of the method to typical cases clearly illustrate the advantage of OBP.

A drawback to DS SS with OBP arises if multiple uplinks are needed. Each uplink signal must have a CDMA receiver including code generator, synchronisation etc. For a large number of user signals, this multiplicity of receivers becomes unwieldy and expensive. By contrast OBP methdos for handling multiple users with FH AJ are well developed and therefore is the obvious way to go. A possible flexible scheme for on-board processing with adjustable processing gain allowing for multiple frequency-hopped uplink is described in Section 8.2.

6.1.5 LPE Techniques

6.1.5.1 Introduction

Low probability of exploitation (LPE) is a term describing an attempt to reduce the ability of an enemy to exploit signals transmitted by a communicator. Forms of exploitation range from simple detection through interception, direction finding, traffic anlysis, to message deciphering. Traffic anlysis is thwarted, especially for downlink SATCOM, by cover signals. Message deciphering is overcome by proper encryption. The threat of detection, interception, and direction finding are countered by low probability of detection (LPD) or low probability of interception (LPI) techniques. There is mixed use of the terms LPE, LPI, and LPD. Here LPE is the broadest term and is meant to encompass LPD and LPI.

In this subsection, only LPD is considered since without detection there can be no exploitation of any of the other forms. Furthermore, the reduction in detectivity of only the uplink signals from terminals is considered since a covert downlink is not feasible.

As for AJ ECCM , LPD can be implemented with waveform techniques and with antenna beam control. Similarly, these two techniques can be combined to give a reduction in detectivity equal, in dB, to the sum of the two individual reductions. Therefore, in this subsection, we will concentrate on the detectability of a number of LPD waveforms by a variety of detection types. If the resulting detection range is too large, then beam control would be required.

Below a very brief description of beam control techniques is given first. The remainder of this LPE subsection considers LPD implemented via SS techniques. First a general detector is described and a measure of performance for the general detector is given. Then the performance of specific combinations of waveform and detector are given. Finally a general discussion of uplink detection is presented.

6.1.5.2 Beam Control For LPD

Beam control for terminals consists primarily of using narrow beam widths with low sidelobes. A rule of thumb for antenna patterns is that sidelobe levels cannot reliably be suppressed below 0 dB_i. In other words, if the peak gain of the antenna is G dB_i, then some sidelobes can be expected at no less than O dB_i. Considerable work has gone into low sidelobe antennas. Extraordinary measures such as using lenses, and absorbing shrouds have been found to reduce sidelobes below O dB_i. Unfortunately, even such extraordinary measures tend to degrade if the antenna is installed on metalic structures such as ships and aircraft. Thus, the O dB_i tends be the best one can expect in practice.

To achieve narrow, low sidelobe antennas, it is clear from the O dB_i rule that higher frequency bands such as EHF are preferred. Lasercom provids even better LPD because of its narrow beams with very low sidelobes.

For ground based terminals, another technique related to antenna patterns becomes available in the form of terrain shielding. By careful selection of a terminal site relative to hills, buildings, trees, etc., there can be a considerable reduction of sidelobe radiation in the direction of would-be interceptors. Terrain shielding is not available to ship-borne and airborne terminals.

6.1.5.3 Performance Measure For The General SS Detector

Spread spectrum (SS) waveforms are employed in military systems primarily for their anti-jam capability. A side benefit of the SS waveforms is that they are more diffucult to detect by interceptors than unspread signals. Although there is a body of literature on the detectability of SS waveforms, there appears to be no universally accepted measure of detectability performance. One such performance measure is provided below.

A general form of the interceptor detector for SS waveforms is shown in Fig 6.8. The signal plus system noise from the interceptor's antenna are fed to bandpass filter of center frequency f_c and bandwidth W. The filtered signal plus noise then passes through a nonlinearity chosen to reveal various spectral lines of the signal. If the spectral line of interest is at some offset frequency f_o, then a mixer is used to bring the line plus the noise at f_o down to baseband. The baseband signal is integrated over a time T to generate a sample value v which is compared to a preset threshold, v_t. If $v>v_t$, a detection is declared whereas if $v<v_t$, it is decided that no signal was received. Then, the probability of false alarm, detection P_d, is the probability $v>v_t$ when a signal is present and the probability of false alarm, P_f, is the probability that $v>v_t$ when no signal is present.

It was desired to find some representation of P_d and P_f as a function of input signal and noise. Numerous possibilities exit, but the one chosen here was to plot P_d on a linear scale between

O and 1 against input S/N_o in dB-Hz or against SNR in dB for a fixed P_f. Here, S is the average signal power and $N_o/2$ is the two-sided power spectral density of the noise. The particular form of SNR will depend upon the application and could be E_c/N_o for DS, E_h/N_o for FH, or E_b/N_o. Here, E_c, E_h, and E_b are energy per chip, hop and bit, respectively. This form of representing performance has been used elsewhere such as in [6.6].

After calculating this form of performance curve for a wide range of combinations of waveforms and detectors, it was found that certain charactaristics always pertain. A typical curve is shown in Fig. 6.9. At lower values of S/N_o (or SNR), the P_d approaches the fixed value of P_f. At higher values of S/N_o, P_d approaches unity. There is a transition region between these two extremes that is only a few dB-Hz (or dB for SNR) wide. The value P_d=0.5 will arbitrarily be taken as a reference point and the corresponding value of S/N_o or SNR will be called the transtion or cross-over value, $S/N_o|_c$ or SNRc. The value of $S/N_o|_c$ becomes an important measure of the performance of an interceptor because, if the intercepted S/N_o is a dB-Hz or so below $S/N_o|_c$ then $P_d \approx P_f$ and the signal will never be detected properly. Conversely, if S/N_o is a dB-Hz or so above $S/N_o|_c$, then $P_d = 0.99 \approx 1.0$ and the interceptor has a very good probability of detecting the signal.

Once $S/N_o|_c$ is determined, then it is straight forward to calculate the detection range from the geometry and parameters of a particular transmitter-interceptor combination. Similarly, a comparison of the susceptability of detection of various SS waveforms by various detector types can be made easily by calculating the values of $S/N_o|_c$. Yet another application would be in comparing the performance of interceptors to the performance of authorized receivers. Here, the more appropriate value would be SNR, such as E_h/N_o for FH, since it is known what value of SNR is required to achieve a desired P_b.

6.1.5.4 Performance Of Various Combinations Of Detectors and SS Waveforms

In table 6.2 are listed various types of detectors that can be used for various waveforms [6.6][6.13]. The wideband radiometer, or energy detector, is the classical form of the general detector shown in Fig. 6.8. It is usually the standard against which the other forms are compared. The nonlinearity is a squaring and no offset frequency is used. Thus it is an energy detector and can therefore detect any signal with sufficient SNR. This universality is also a problem since other signals within the band W will give false alarms.

For FH waveforms, a specialized detector is the filter bank combiner wherein a bank of radiometers are used each with a band W equal to the individual hop width which for fast FH is R_h. The idea is to provide a large SNR in the filter that contains the hop. In practice, since W_{ss} is so large, the number of filters and radiometers is kept to a small fraction of that required to cover all of W_{ss}. A good method of implementing the reduced filter bank is with a chirp transform implemented with SAW's (See Sec. 8.2). In the limit, a lab spectrum analyser set to a resolution of about R_h is equivalent to a single filter/radiometer and picks up hops that, at random, land in the resolution bandwidth.

DS waveforms are thought, erroneously, to provide better LPD capability than FH because in the frequency plane the DS signal can fall below the noise floor whereas the FH signal can pop up randomly above the noise floor. However, a number of nonlinearities can alter the DS waveform to have strong easily detected spectral lines. In the squaring carrier detector, a square law generates a spectral line at $2f_c$. In the delay-and-multiply chip-rate detector, the nonlinearity is obtained by multiplying the input by a delayed version of itself where the delay is approximately half of a chip period. A strong spectral chip rate is produced.

It has been found for a wide variety of combinations of SS waveforms and detector types that the S/N_o required to achieve a given P_d and P_f is well approximated by

$$S/N_o = c\eta\sqrt{W/T}\,\left[Q^{-1}(P_f) - Q^{-1}(P_d)\right] = c\eta\sqrt{W/T}\;d \qquad (1)$$

where c is a constant depending upon the SS waveform and detector configuration and d is a function of P_f and P_d and is some times called the "detectivity". The functions Q^{-1} are the inverse of

$$Q(z) = \frac{1}{2\pi}\int_z^{\infty} e^{-u^2/2}\,du. \qquad (2)$$

The correction factor, η, corrects for the fact that the probability distribution out of the nonlinearity is not gaussian. However, under most cases of practical interest, the product TW is large which results in a good gaussian approximation and usually, η is very close to unity. Eqn (1) describes the S curve in the performance curve of Fig. 6.8.

The constant, c, is now a key value. it provides a means of comparing all combinations of waveforms and detectors. Some of these values are listed in Table 6.2. For some of the combinations, c has not been determined. For the filter band combiner, W in (1) is R_h and the P_d is the P_d for the single channel containing the hop. To overall P_f is calculated by combinatorial methods.

As examples of the performance curves described by (1), consider W=10MHz, T=1s, η=1 (for large TW), and c=1. In Figure 6.10 are plotted 2 curves; one for P_f=0.01 and one for P_f=0.0001. For other values of c, the curves are shifted to the left by 10 log c dB-Hz.These curves for c=1 describe the performance of a wideband radiometer detecting an FH signal with a hopping band of 10 MHz or detecting a DS signal with a 5-MHz chiprate (actually c=1.1), or detecting a FH/DS hybrid spread over 10 MHz. It also describes the performance of a squaring carrier detector detecting a 5 Mchip/s DS.

It is also useful to plot the curves as a function of SNR. As an example, consider a data bit rate of 1 kb/s either spread by the above FH or DS examples. E_b/N_o of 10 to 12 dB is all that these detectors require to obtain good P_b. However, the legitimate receiver also needs about this much SNR to achieve adequate P_b. Thus, for this example, the interceptor need be no closer to the transmitter than the legitimate receiver.

It is concluded from the above that for a given spread band, that one SS waveform is as detectable as another. Since the required S/N_o is proportional to W the best strategy of the communicator is to make W_{ss} as wide as possible regardless of waveform. Therefore, FH is the preferred waveform since the widest W_{ss} is achievable in practice with FH. The addition of DS to FH has erroneously been thoght to decrease detectivity. In fact, it doesn't change the detectivity against a radiometer at all, but may decrease slightly the performance of some of the specialized detectors.

6.1.5.5 Discussion Of Detection Uplink Signals

Unfortunately, against the detectors described above, SS uplink signals are usually easily detectable for two reasons. First, the reduction in detectability recreases only as $1/\sqrt{W}$ and this decrease can be more than offset by the interceptor by integration time T so that detectivity goes up by $\sqrt{T}$. The value of T can be as long as the message length. Conversely, the intended receiver must receive enough power that integration over a symbol period yields a correct bit and generally the symbol period is much less than T.

The second advantage of an uplink interceptor is that the interceptor is generally much closer to the transmitter than the satellite so that the interceptor suffers from far less space loss. For example at 44 GHz, the space loss to a geosynchronous satellite is approximately 215 dB whereas an interceptor at 100 km has a loss of only 165 dB so that the interceptor enjoys a 50 dB advantage. Consider a 1-m dish with a gain of about 50 dB_i. Since it was seen that sidelobes typically are no more than O dB_i an interceptor in the sidelobe is still intercepting a signal with as much power as received by the intended satellite. Thus, the space loss difference cancels the advantage of the sidelobes.

Covert uplink transmissions are very difficult to achieve especially in a marine environment. The recommendation is to spread as much as possible, keep the message length as small as possible, and use an antenna with as low sidelobes as possible. Note that burst communications actually increases the detectability of intercepted signals because of the increased power levels. However, if the interceptor is using spatial or frequency scanning, a burst might be missed. A ship's captain would be very uncomfortable with having to depend on the chance that the interceptor is looking another way when a burst is sent.

6.1.6 ECCM Performance Of A Hypothetical EHF SATCOM System For NATO

6.1.6.1 System Characteristics

One geostationary satellite situated over the east Atlantic can provide sufficient coverage for communication among terminals within the NATO ACE (and also Atlantic) region. To provide coverage at lattitudes above approximately 65° (especially if communications from the polar regions are required), a constellation of satellites utilizing inclined orbits (e.g. 63° inclined 20000 km, 12 hour orbits) will be required. Inter-satellite links may be used to provide connectivity between users accessing different satellites.(See Chapter 10)

The EHF satellites serving the mobile military users are expected to use the 44 GHz uplink and 20 GHz downlik frequency bands with the satellite bandwidth available in the uplink and the downlik directions being 2 GHz and 1 GHz respectively. Frequency hopping is expected to be used as the spread-spectrum AJ modulation technique so as to fully exploit the available tranmission bandwidths (and also to minimize the disturbances from high altitude nuclear bursts).

On-board processing involving dehopping/rehopping or dehopping/demodulation/remodulation/rehopping techniques are expected to be utilized in these satellites. Such a processing transponder will provide AJ performance improvement over that can be provided by a conventional non-processing transponder. Furthermore, such a processing transponder will transform the available 2 GHz uplink bandwidth into a GHz downlink bandwidth, hence permitting the full utiliza of the wider spreading bandwidth available in the uplink direction (See Chapter 11).

It is assumed that these satellites will use multibeam receive antennas with adaptive spatial nulling capability and multiple spotbeam transmit antennas for increased jamming resistance.

6.1.6.2 Jammed Traffic Capacity

The satellite is assumed to dehop, regenerate and rehop the received uplink frequency hopping signals before retransmission back to earth. With such a transponder, the total bit error rate of the end-to-end SATCOM link, BER_T, will be approximately equal to the sum of the uplink and downlink bit error rates, BER_U and BER_D:

$$BER_T = BER_U + BER_D - 2BER_U . BER_D \approx BER_U + BER_D \qquad (1)$$

In the presence of uplink jamming, BER_T is approximately equal to BER_U. Similarly for the case of downlink jamming, BER_T will be approximately equal to BER_D.

BER_U and BER_D are a function of the energy-per-bitt-noise-density ratio, E_b/N_o, at satellite receiver and the ground terminal receiver inputs respectively after dehopping. The relation between BER and E_b/N_o depends on the modulation/demodulation and the error correction coding/decoding method used. A frequency hopping modem (transponder) using coded MFSK modulation, noncoherent demodulation and soft decision decoding and delivering less than 1 error in 10^3data bits for $E_b/N_o \le 10$ dB under partial band jamming conditions is within the current state of the art.

a) Uplink Jamming

The total uplink data rate R_{dU} that can be supported by a transmitting SATCOM terminal in the presence of uplink jamming, while maintaining a minimum acceptable uplink E_b/N_o is given by

$$R_{dU} = \frac{1}{M_U . (E_b/N_o)_U} \cdot \frac{P_T}{\dfrac{kT_S L_U}{G_{RS}} + \dfrac{P_{JU}}{\alpha B_{SU}}} \qquad (2)$$

where

P_T = SATCOM terminal EIRP

P_{JU} = uplink jammer EIRP

G_{RS} = satellite receive antenna gain in the SATCOM terminal direction.

α = satellite receive antenna nulling in the jammer direction.

T_S = effective noise temperature of satellite receiver

k = Boltzmans constant

B_{SU} = uplink spreading (hopping) bandwidth

L_U = uplink free space loss

$(E_b/N_o)_U$ = minimum acceptable energy per bit-to-noise density ratio after dehopping at the satellite.

M_U = margin for atmospheric and rain losses at uplink frequency.

In equation (2), the satellite range from the terminal and from the jammer (and hence the uplink free space losses) have been assumed to be equal.

Example 1: Uplink jamming of geosynchronous EHF satellite

Uplink frequency : 44 GHz
Satellite range : 36000 km

P_T = 60 dBW (100 W TWTA + 40 dB antenna gain)

G_{RS} = 20 dB

α = 25 dB

T_S = 100 K°

B_{SU} = 2 GHz

L_U = 216.4 dB

(E_b/N_o) = 10 dB

M_U = 15 dB

P_{JU}, dBW	R_{dU}, bits/sec
90	5240
100	5120
110	4160
120	1450

b) Downlink Jamming

The total data rate that can be supported on the downlink to receiving SATCOM terminals using a downlink satellite EIRP of P_{SAT} in the prence of downlink jamming is given by

$$R_{dD} = \frac{1}{M_D\,(E_b/N_o)_D} \cdot \frac{P_{SAT}}{\dfrac{kT_RL_D}{G_{RT}} + \dfrac{P_{JD}L_D}{\beta\,B_{SD}L_JL_{AJ}}} \quad (3)$$

where

P_{SAT} = downlink satellite EIRP in the receive terminal direction

P_{JD} = jammer EIRP in the receive terminal direction

G_{RT} = receive terminal antenna gain (beam centre)

β = main lobe to side lobe ratio for the receive terminal antenna

T_R = effective noise temperature of the receive terminal.

B_{SD} = downlink spreading (hopping) bandwidth

L = downlink free space loss

L_J = free space loss between jammer and the receive terminal

L_{AJ} = atmospheric and rain losses of the jammer to receive terminal path.

$(E_b/N_o)_D$ = minimum acceptable energy per bit to noise density ratio after dehopping at the terminal.

M_D = margin for atmospheric and rain losses at the downlink frequency.

The jammer is assumed to be jamming the receive terminal at the sidelobes of its receive antenna. The free space path loss between the jammer and the terminal is given by

$$L_J = \left(\frac{4\pi}{c}\, f_D\, r_{JT}\right)^2 \quad (4)$$

where f_D is the jamming (downlink) frequency, r_{JT} is the separation between jammer and the terminal and c is the velocity of light

Example 2: Downlink jamming of an EHF SATCOM terminal.

Downlink frequency : 20 GHz

Satellite range : 36000 km

Jammer to terminal separation : 100 km

P_{SAT} = 50 dBW (spot beam)

G_{RT} = 33 dB

β = 33 dB

T_R = 500° K

B_{SD} = 1 GHz

L_D = 209.6dB

L_J = 158.5dB

L_{AJ} = 6dB

$(E_b/N_o)_D$ = 10dB

M_D = 6 dB

P_{JD}, dBW	R_{dD}, bits/sec
20	8.00×10^5
30	7.96×10^5
40	7.60×10^5
50	5.27×10^5
60	1.29×10^5
70	1.51×10^4
80	1.54×10^3

6.2 INFORMATION CODING

In digital communications coding of the information to be transmitted is usually split into two areas of coding

- the source coding for data compression
- the forward error control/correcting (FEC) coding.

Both classical areas of coding are of fundamental interest in communications. For military communications data compression techniques offer an increase of processing gain of the used spread spectrum system and FEC coding can be exploited to improve the transmission quality in terms of the bit error rates significantly.

For voice and image compression a large number of algorithms have been studied extensively and have been realized in hardware. To name a few methods:

- Continuously variable slope delta modulation (CVSD)
- Differential pulse code modulation (DPCM)
- Linear Predictive Coding (LPC)
- Adaptive predictive coding (APC)
- Transform coding (Fourier, Hadamard...)
- Run-length coding
- Tandem combinations such as Hadamard -DPCM

Currently, methods for combining source and FEC coding are studied with the aim of utilizing the redundancy of the information in a more efficient way.

As source coding also FEC coding has been studied extensively for many years. A wide range of coding methods for all kinds of applications such as for mobile satellite or terrestrial communications, communications in an environment with severe jamming and also for magnetic tape recording, compact discs, fault tolerant memories etc. are available for implementation. It depends on the application as to what coding scheme is selected. Block codes can be soft-decision decoded using the Viterbi algorithm. Since soft-decision decoding results in an

additional coding gain of about 2dB(Gaussian channel) especially convolutional codes are attractive in the case where soft-decision values are available for decoding. The most common convolutional code currently used might be the constraint length K=7, rate 1/2 code. This power-efficient code requires a 64-state Viterbi decoder, which can be realized by state-of-the art technology on a single chip. The class of Reed-Solomon (RS) codes may be regarded as the most common block codes, implemented e.g. in compact disc systems. A very efficient coding scheme is formed by the combination of RS and convolutional codes (concatenated coding). A number of studies have shown, that with such schemes system degradations due to worst case jamming can be almost compensated. To evaluate the status of information coding at the end of this century it is worth identifying the areas of current research.

- Channel adaptive coding with variable rate.
  In time-varying channels for good channel conditions less efficient coding is required than for bad channel conditions. So, the throughput can be increased by adapting the coding rate to the momentary channel state. Variable rate coding is achieved with code "puncturing" techniques.

- Coded Modulation
  Since Ungerboeck introduced the so-called trellis codes, which are constructed with recpect to the used modulation (to the distance properties in the euclidean signal space), many new codes have been propesed offering anew class of power - and bandwidth efficient coding/modulation schemes.

- Applications to high data rate communications.
  The application of FEC coding in high-data-rate communications is determined by the speed of decoding. Currently single chip RS-decoders can be realized with a speed of about 100 Mbit/sec. Single chip realizations of 64-state Viterbi decoders are commercially available with a decoding speed of about 25 Mbit/sec. Utilizing new technology such as systolic arrays single chip Viterbi decoders running at a speed of about 100 Mbit/sec can realized in the near term. Another approach for high speed decoding might be to employ concatenated coding schemes consisiting of lower complex basic codes in a parallel structure. It turns out, that this approach promises single chip decoders with several hundred Mbit/sec decoding speed realized with available technology.

- Generalized Concatenated Coding
  The concept of code concatenation, introduced by D. Forney, was generalized. The generalized coding construction offers the possibility to form very long codes still decodable, since decoding can be partitioned in several steps.

The above may illustrate that information coding techniques are already highly advanced so that at the turn of this century coding can be regarded as a well-known, standard technique used in almost all areas of communications

Appendix 6B discusses the efficiency of FEC techniques for the following channels:

a) Channel with white Gauisian noise (AWGNC)
b) Rayleigh -fading channel (mobile satellite coms)
c) DS spread-spectrum in a channel with jamming
d) FH spread-spectrum in a channel with jamming

6.3 SPEECH CODING

6.3.1 Waveform Coding

Speech compression systems can generally be classified as either waveform coders or Vocoders (i.e., voice coders or analysis-synthesis telephony). These two classes cover the whole range of compressibility from 64000 down to a few hundred bits per second. The important factors which need to be taken into account when comparing different encoding tehcniques are the speech quality achievable in the presence of both transmission errors and acoustic noise, the data rate required for transmission, the delay introduced by processing, the physical size of the equipment and the cost of implementation (a function of coder complexity which can be measured by the number of multiply-add operations required to code speec. usually expressed in millions of instructions per second "MIPS").

The most basic type of waveform coding is pulse code modulation (PCM) consisting of sampling (usually at 8 kHz), quantising to a finite number of levels, and binary encoding. The quantiser can have either uniform or non-uniform steps giving rise to linear and logarithmic PCM respectivly. Log-PCM has a much wider dynamic range than linear PCM for a given number of bits per sample, because low amplitude signals are better represented, and as a result logarithmic quantisation is nearly always used in wideband speech communications applications. A data rate of 56 to 64 kbit/s is required for commercial quality speech and lower rates for military tactical quality.

There are many variations on the basic PCM idea, the most common being differential encoding and adaptive quantisation. Each variation has the object of reducing the rate required for a given speech quality, a saving of approximately 1 bit per sample (8 kbit/s) being achieved when each is optimally employed. In differential PCM (DPCM) the sampled speech signal is compared with a locally decoded version of the previous sample prior to quantisation so that the transmitted signal is the quantised difference between samples In adaptive PCM (APCM) the quantiser gain is adjusted to the prevailing signal amplitude, either on a short term basis or syllabically. By controlling the adaption logic from the quantiser output, the quantiser gain can be recoreved at the receiver without the need for additional information to be transmitted. Adaptive differential PCM (ADPCM) is a combination of DPCM and APCM which saves 2 to 4 bits per sample compared with PCM, thus giving 48 to 32 kb/s with high quality speech.

It is interesting to note that although the principle of DPCM has been known for 30 years, it was not possible to standardise such a 32 kb/s coder until 1983 [6.14], after efficient and robust algorithms became available. These adaptive algortihms are efficient in the sense that they adapt quantisation and prediction synchronously at the encoder and decoder without transmitting explicit adaption information. They are robust in the sense that they function reasonably well even in moderate bit-error environment.

There is another adaptive approach to producing high quality and lower bit-rate coder which is called "adaptive subband coding" which divides the speech band into four or more contiguous bands by a bank of filters and codes each band using APCM. After lowering the sampling rates in each band, an overall bit rate can be obtained while maintaining speech quality; by reducing the bits/sample in less perceptually important high-frequency bands. Bands with low energy use small step sizes, producing less quantisation noise than with less flexible systems. Futrhermore, noise from one band does not affect other frequency bands. Coders operating at 16 kb/s using this technique have been shown to give high quality but with high complexity [6.15].

When the number of quantisation levels in DPCM is reduced to two, delta modulation (DM) results. The sampling frequency in

this case is equal to the data rate, but it has to be well above the Nyquist frequency to ensure that the binary quantisation of the difference signal does not produce excessive quantisation noise. Just as with PCM, there are many variations of DM, and the right hand side of Fig o illustrates some of them. The most important form of DM used in digital speech communications is syllabically companded DM; there are a number of closely related versions of this, examples being continuously variable slope DM (CVSD) and digitally controlled DM (DCDM). The data rate requirements are a minimum of about 16 kbit/s for military tactical quality speech and about 48 kbit/s for commercial quality.

When operated at data rates of 12 kbit/s and lower, the speech quality obtained with PCM and DM coders is poor, and consequently they cannot be used as narrow band devices. However, the principles of operation of wideband coders are useful in analysis-synthesis telepnony once significant redundancy has been removed from the speech waveform.

6.3.2 Analysis-Synthesis Telephony

Analysis-synthesis telephony techniques are based on a model of speech production:

There are two basic types of speech sound which can be produced, namely voiced and unvoiced sounds. Voiced sounds occur when the vocal cords are tightened in such a way that the subglottal air pressure forces them to open and close quasi-periodically, thereby generating "puffs" of air which acoustically excite the vocal cavities. The pitch of voiced sounds is simply the frequency at which the vocal chords vibrate. On the other hand, unvoiced sounds are produce by forced air turbulence at a point of constriction in the vocal tract, giving rise to a noise-like excitation, or "hiss".

In channel vocoding the speech is analysed by processing through a bank of parallel band-pass filters, and the speech amplitude in each frequency band is digitized using PCM techniques. For synthesis, the vocal and nasal tracts are represented by a set of controlled gain, lossy resonators, and either pulses or white noise are used to excite them. In pitchexcited vocoders it is derived by non-linear processing of the speech signal in a few of the low frequency channels combined into one. Pitch-excited vocoders require data rates in the range from 1200 into 2400 bit/s and yield poor quality speech, whereas voice-excited vocoders will provide reasonable speech quality at 4800 bit/s and good quality at 9600 bit/s.

A formant vocoder is similar to channel vocoder, but has the fixed filters replaced by formant tracking filters. The centre frequencies of these filters along with the corresponding speech formant amplitudes are the transmitted parameters. The main problem is in acquiring and maintaining lock on the relevant spectral peaks during vowel-consonant-vowel transitions, and also during periods where the formants become ill-defined. The data rate required for formant vocoders can be as low as 600 bit/s, but the speech quality is poor. The minimum data rate required to achieve good quality is 1200 bit/s, but to date this result has only been obtained using semi-automated analysis with manually interpolated and corrected formant tracks.

The third method of anlysis-synthesis telephony to have achieved importance is linear predictive coding. In this technique the parameters of a linearised speech production model are estimated using mean-square error minimisation procedures. The parameters estimated are not acoustic ones as in channel and formant vocoders, but articulatory ones related to the shape of the vocal tract. For a given speech quality, a transmission data rate reduction in comparison with acoustic parameter vocoding should be achieved because of the lower redundancy present. Just as with channel and formant vocoders, excitation for the synthesizer has to be derived from a separate eanalysis, the usual terminology being pitch-excited or residual excited, corresponding to pitch or voice excitation in a channel vocoder.

LPC is a very active area of speech research, and new results appear regularly. At present data rates as low as 2400 bit/s have been achieved for pitch-excited LPC with reasonable quality speech, and in the range from 8 kbit/s to 16 kbit/s for residual excited LPC with good speech quality.

The application of vector quantisation (VQ), a fairly new direction in source coding, has allowed LPC rates to be dramatically reduced to 800 b/s with very slight reduction in quality, and further compressed to rates as low as 150 b/s while retaninig intelligibility [6.16, 6.17]. This technique consists of coding each set or vector of the LPC parameters as group instead of individually as in scalar quantisation. Vector quantisation can be used also for waveform coding.

A good candidate for coding at 8 kb/s is multipulse linear predictive coding, in which a suitable number of pulses are supplied as the excitation sequence for a speech segment-perhaps 10 pulses for a 10-ms segment. The amplitudes and locations of the pulses are optimised, pulse by pulse, in a closed-loop search. The bit rate reserved for the excitation information is more than half the total bit rate of 8 kb/s. This does not leave much for the linear predictive filtre information, but with VQ the coding of the predictive parameters can be made accurate enough.

For 4 kb/s coding, code excited or stochastically excited linear predictive coding is promising. The coder stores a repertory of candidate excitations, each a stochastic, or random sequence of pulses. The best sequence is selected by a closedloop search. Vector quantization in the linear predictive filter is almost a necessity here to guarantee that enough bits are available for the excitation and prediction parameters. Vector quantization ensures good quality by allowing enough candidates in the excitation and filter codebooks.

Table 6.3 below compares tradeoffs for representative types of speech coding algorithms [6.18]. It shows the best overall match between complexity, bit rate and quality. A coder type is not necessarily limited to the bit rate stated. For example, the medium -complexity adaptive differential pulse-code modulation coder can be redesigned to give communication-quality speech at 16 kb/s instead of high-quality speech at 32 kb/s. In fact, a highly complex version can provide high-quality speech at the lower bit rate. Similarly lower-complexity multipulse linear predictive coding can yield high-quality coding at 16 kb/s, and a lower-complexity stochastically excited linear predictive coder (LPC) can be designed if the bit rate can be 8 kb/s instead of 4 kb/s.

Cost is also tradeoff factor, but it is hard to quantify in a table. The cost of coding hardware generally increases with complexity. However, advances in signal processor technology tend to decrease cost for a given level of complexity and, more significantly, to reduce the cost difference between low-complexity and high-complexity techniques.

Of course, as encoding and decoding algorithms become more complex they take longer to perform. Complex algortihms introduce delays between the time the speaker utters a sound and the time a coded version of it enters the transmisison system. These coding delays can be objectionable in two-way telephone conversations, especilaly when they are added to delays in the transmission network and combined with uncanceled echoes. Coding delay is not a problem if the coder is used in only one stage of coding and decoding, such as in voice storage. If the delay is objectionable because of uncanceled echoes the addition of an echo canceler to the voice coder can eliminate or mitigate the probelems. Finally, coding delay is not a concern if the speech is merely stored in digital form for later delivery.

6.3.3 Conclusions

Speech communication is and will remain in the foreseeable future the main mode of communication, not only for civil but also for strategic/tactical military applications. Digital speech processing is, consequently, an essential ingredient of the evolving ISDN's to be used by both civil and military users. A fully implemented ISDN is seen as a real asset to national security and preparedness. End-to-end digital communications of the kind promised by ISDN are well suited to secure communications. Furthermore, the ubiquity, connectivity, and interoperability inherent in the concept will be most valuable in emergency situations requiring reconfigured communications.

Speech coding methods have been standardised internationally at 64 kb/s (PCM) and 32 kb/s (ADPCM) and coders at these rates are being used in the common-user switched telephone networks. Continuously Variable Slope Delta Modulation (CVSD) has also been standardised in NATO for tactical military communications. There are also both civil and military requirements for speech coders operating at speeds of 16kb/s and below e.g., for mobile land and maritime communications. For HF communications and LOS radio and satellite communications under heavay jamming, vocoders operating at 2.4 kb/s and even below are required. Secure voice using 4 kHz nominal analogue channels also requires speech coders operating at speeds of 4.8 kb/s and below. Speech coding is also required for high-fidelity voice (HFV) with 7 and 15 kHz bandwidth as well as for Digital Circuit Multiplication and for longer-term applications, i.e., in the evolving broadband ISDN when "Asynchronous Transfer Mode" (ATM) of operation will be implemented.

It is to be noted that there are important operational requirements in NATO for interoperability between systems using different speech coders; this necessitates standardisation and agreements on interfaces/gateways where code, rate and other (signalling, numbering) conversion take place.

To achieve good quality below 32 kb/s codes must take increasing advantage of the constraints of speech production and perception. At transmission rates below 16 kb/s quality diminishes significantly, requiring more of the, as yet, poorly known properties of speech production and perception. Also at the lower transmission rates, the computational complexity to implement the coding algorithms increases, while the ability to handle nonspeech-like sounds-such as music and voice-band data diminishes. Typically too, the encoding delay increases as the transmission bit rate decreases.

The primary challenge, then is to develop new understanding that will significantly elevate the speech-quality curve for the lower bit rates, even with substantial but acceptable increase in complexity.

The research frontier in coding currently centers on ways to achieve good quality at transmission rates of 9.6 kb/s and below. Undoubtedly, increased computational complexity will be required to elevate the quality of low bit-rate codes, which must extensively use the known redundancies of speech production and perception. Breakthroughs will occur only when new properties of redundancy are found [6.19].

It is expected that in the timeframe considered in this report there will be 8 kb/s or even lower-rate codecs available for use in SATCOM with qualities comparable to that of 64 kb/s PCM voice.

It is interesting to note that the VQ technique which is so useful for reducing rates for speech signals can also be used for any waveform coding including the coding of images, a subject which has not been treated in this report but where important advances in data compression are taking place.

A complete treatment of speech processing techniques can be found in [6.20].

6.4 SATELLITE HARDENING

The following are measures that may be taken to harden satellites against kinetic energy weapons [6.12].

6.4.1 Fragmentation Shields

Two types of fragmentation shields have been investigated:

a) Sacrificial shields,
b) Bumper shields,

The impact of shielding type (a) above from frontal attack alone on the all up mass of the satellite is an increase of 4000 Kg for a base satellite mass of 2600 Kg. Using bumper shields of type (b) the mass impact would be reduced by 1000 Kg.

The actual performance of shielding is of course strongly dependent on the assumed threat and particle shape. Conical, rod shaped and sphere are most effective in the order given.

However it is now assumed that Kinetic Energy Weapons (KEW's) are credible and will be available in the period 2000/2030.

6.4.2 Laser Hard Antennas

Various material constructions have been investigated for survival against laser attack.

6.4.3 Laser Hard Thermal Blankets

Various material constructions have been investigated including the following: Metal Foils of tantalum, titanium, nickel, *molybdenum* as well as non-metals such as quartz, glasses and ceramics.

6.4.4 Conclusions

The following conclusions are drawn based on the above and the discussions in the Group:

a) Space and ground based kinetic energy weapons are credible.

b) Shields will have to be transparent to RF (a few dB loss) and this will effect the satellite design (Geotto satellite had protection).

c) Thermal blankets can be effective against laser. Solar cells are vulnerable.

d) Laser weapons can be ground or space based. The threat is specified in terms of illumination per cm^2 (W/cm^2) (sun flux is 240 mW/cm^2).The ground-based laser can be more effective on satellites at lower rather than higher orbits; effect is negligible for geosynchronous satellites. Sensors may be easier to damage than melting the satellite material.

e) Satelittes in elliptical orbits, because of their lower altitudes in the southern hemisphere, would be vulnerable to laser attacks from the southern hemisphere. One should pay attention to this kind of attack.

f) The only effective protection is to design the satellite to be capable of coping with the incident laser flux and to provide protection for the sensors.

g) Space-borne lasers can be very effective, space-borne jammers can also be used. While the atomic powered laser can destroy a satellite, the high power jammer (using gyrotrons) can burn out the front-end of the satellite receiver and to use

anti-radar missile.

h)ASAT weapons can be launched from the ground, ships and aircraft. There would be no protection against a direct hit by a homing weapon. Proliferation of satellites and maneuvering them (this needs a warning system) are measures which may be taken to counter ASAT weapons.

6.5 REFERENCES

[6.1] M.K. Simon, J.K. Omura, R.A. Scholtz, and B.K. Levitt, "Spread Spectrum Communications", vols. 1,2, and 3, Computer Science Press, Rockville MD 1985.

[6.2] D.J. Torrieri, "Principles of Military Communications Systems", Artech House, Dedham MA, 1981.

[6.3] D.L. Nicholson, "Spread Spectrum Signal Design: LPE and AJ Systems", Computer Science Press, Rockville MD, 1988.

[6.4] J.S. Bird and E.B. Felstead, "Antijam performance of fast frequency-hopped Mary NCFSK -an overview", IEEE J. Selected Areas Comm, vol SAC-4, pp 216-233, March 1986.

[6.5] P. Tardif and E.B. Felstead, "Measurement of diversity-combining performance enhancement for fast FH", IEEE Milcom 89, Conf Record pp. 16-19, Boston, October, 1989.

[6.6] G.R. Cooper, "Detection of frequency-hop signals", IEEE Milcom '86 Conf. Rec., pp. 10.2.1-10.2.5, October 1986. IEEE Catalog No. CH2323-4/86.

[6.7] H. Urkowitz, "Energy detection of unknown determinictic signals". Proc. IEEE, vol. 55, pp. 523-531, April 1967.

[6.8] D.J. Torrieri, "Principles of Secure Communication Systems". Artech House, Dedham MA (1985), Chap. 4, Interception"

[6.9] D.E. Reed and M.A. Wickert, "Minimization of baud rate spectral lines by delay and multiply receivers,"IEEE Milcom 86 Conf Rec. pp.10.3.1-10.3.5, October 1986 IEEE Catalog No CH2323-4/86.

[6.10] J.D Edell ,"Wideband,noncoherent,frenquency hopped waveforms and their hybrids in low probability-of-intercept communications",Naval Research Laboratory,Washington DC. Tecnical Report No.8025,November 8,1976.

[6.11] M.K Simon,J.K. Omura,R.A. Sholtz,and B.D. Levitt,"Spread Spectrum Communications", vol. 3. Computer Science Press Rockville MD,(1985).

[6.12] R.A.Dillard, "Detectability of spread-spectum signals,"IEEE Tras. Aerospace and Elect.Sys.,vol.AES-15,pp.526-537,.July 1979.

[6.13] E.W.Chander and G.C Cooper "Development and evaluation of an LPI figure of merit for direct sequence and frequency-hop system",IEEE Milcom'85 Conf.Rec.,pp.603-608,October 1985.IEEE Catalog No.85CH2202-0.

[6.14] "CCITT Red Book,Vol III",Geneva,ITU Press.1948.

[6.15] Crochiere R.E.,et al.,"Real-Time Speech Coding" IEEE Trans.On Commun., Vol COM-30.April 1982

[6.16] Buze A., et al., "Speech Coding Based upon Vector Quantisation",.IEEE Trans.Acoust.,Speech and Signal Process.,ASSP-28,No.5,Oct.1980

[6.17] Roucos S.,et al.,"Vector Quanzation for Very-Low-Rate Coding of Speech", Conf.Rec.,1982 IEEE Global Coms Conf.,FL, Nov.29-Dec.2,1982.

[6.18] Jayant N.S., "Coding Speech at Low Bit Rates",IEEE Spectrum,Aug.1986.

[6.19] "Speech Processing Technology",ATT Technical Journal, Sept.Oct.1986 Vol 65, Issue 5.

[6.20] Ince A.N. et al."Speech Analysis and Synthesis and Man - Machine Speech Communications for Air Operations", AGARD Lecture Series, No:170, 1990.

[6.21] Lambert M., Slachmuylders.,"Spacecraft Shielding Against Orbital Debris",41st Congress of the International Astronautical Federation, Oct 6-12,1990,Dresden.

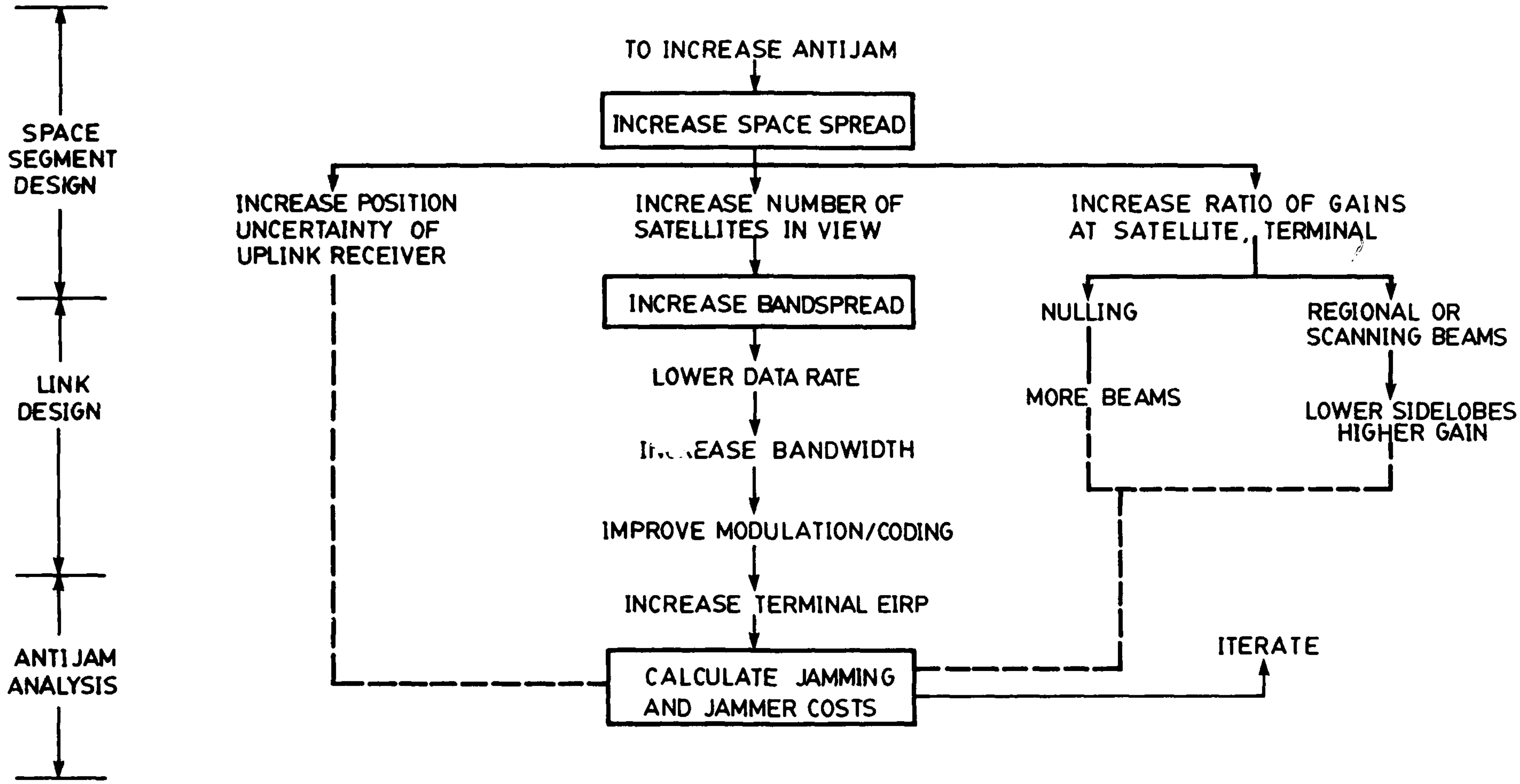


Fig.6.1 Iterative design process for antijam satellite system

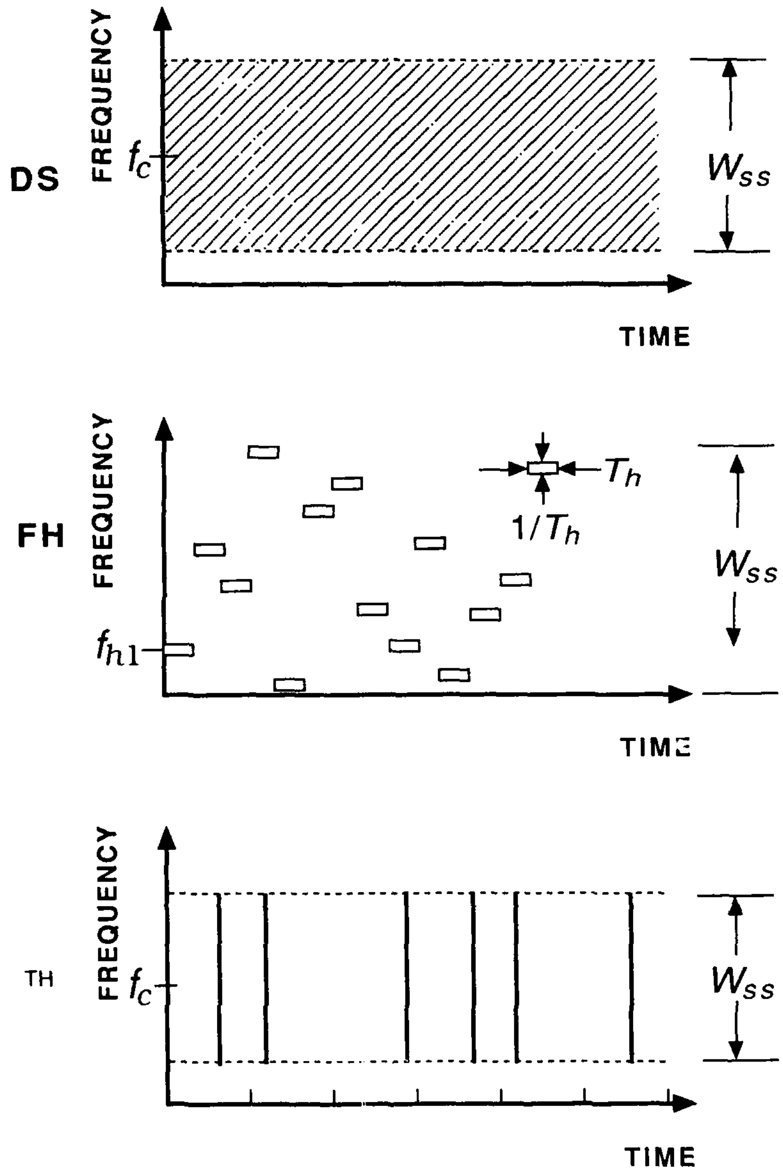


Fig.6.2 A time-frequency representation of DS, FH and TH waveforms

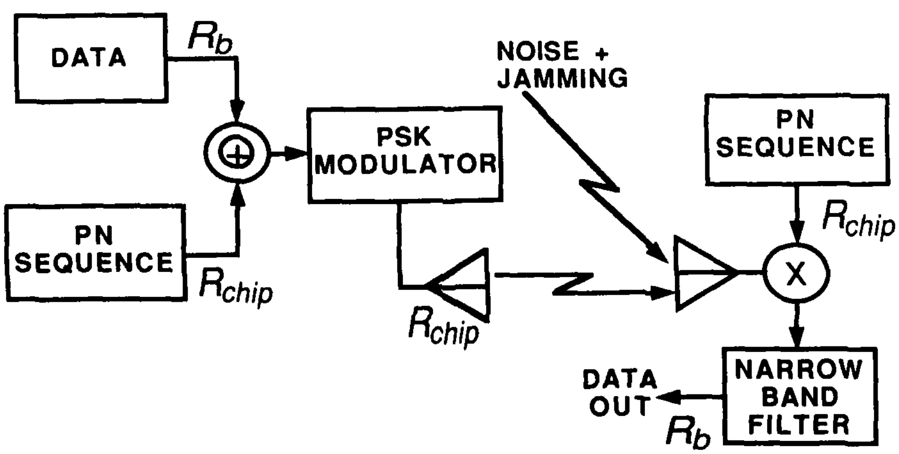


Fig.6.3 A simplified block diagram of a typical DS SS system

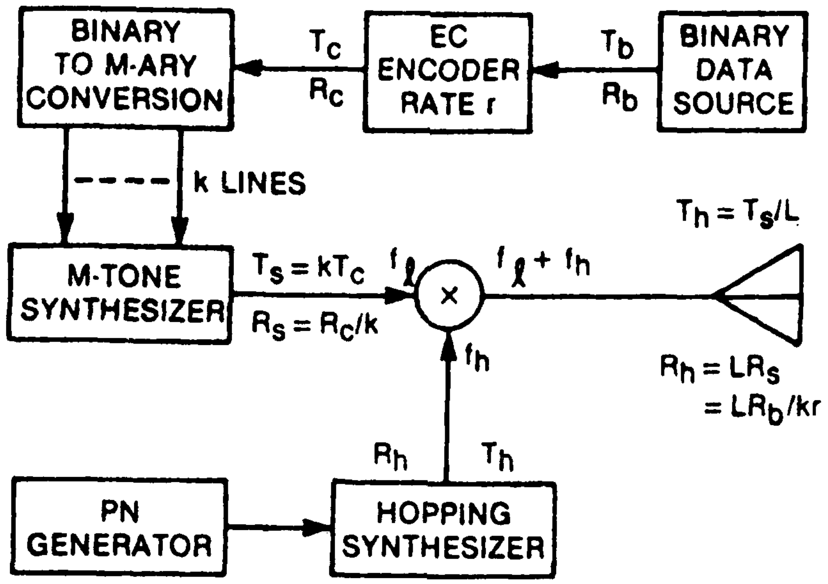


Fig.6.4 Block diagram for a fast hopped M-ary NCFSK transmitter

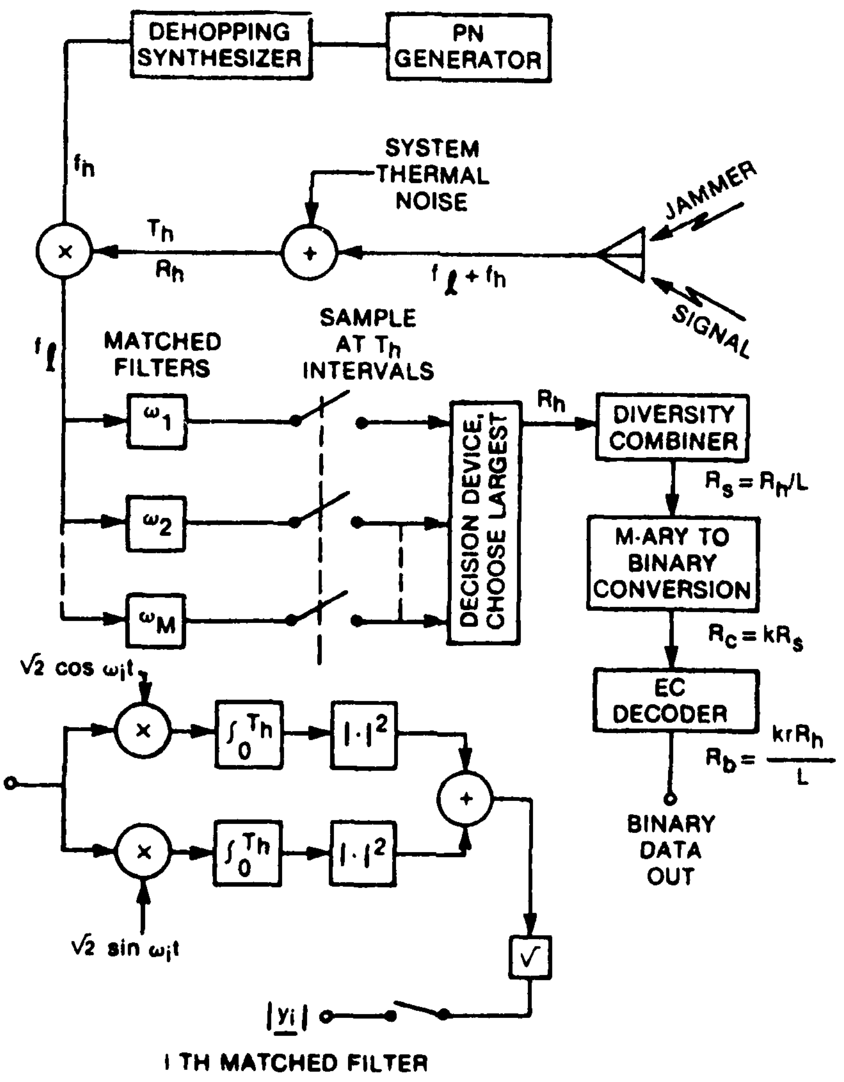


Fig.6.5 Block diagram for receiver for fast hopped M-ary NCFSK

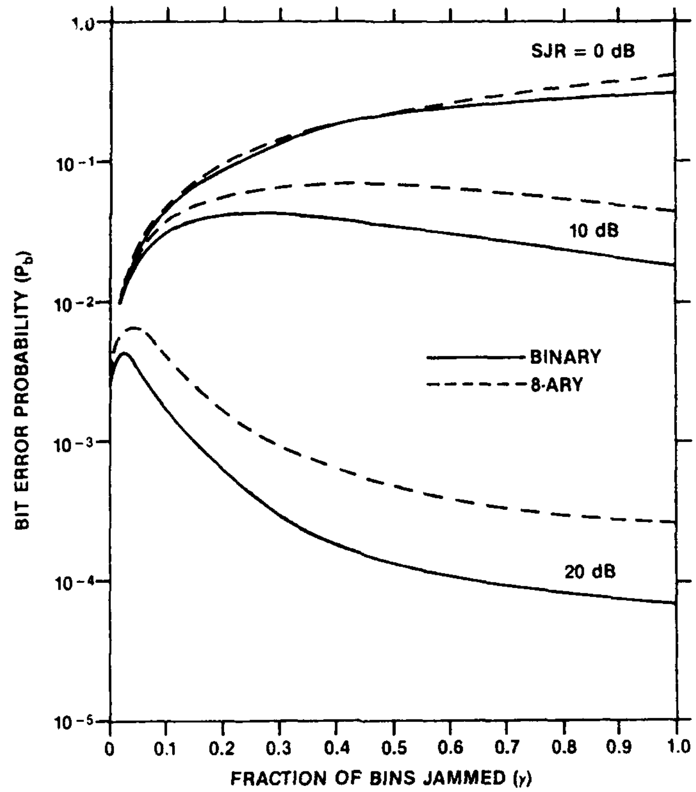


Fig.6.6 P_b versus γ for SNR = 13.35 dB under partial-band noise jamming. P_b as a function of γ for 8-ary NCFSK in the presence of PBN jamming without and with system noise where SNR = 13.35 dB.

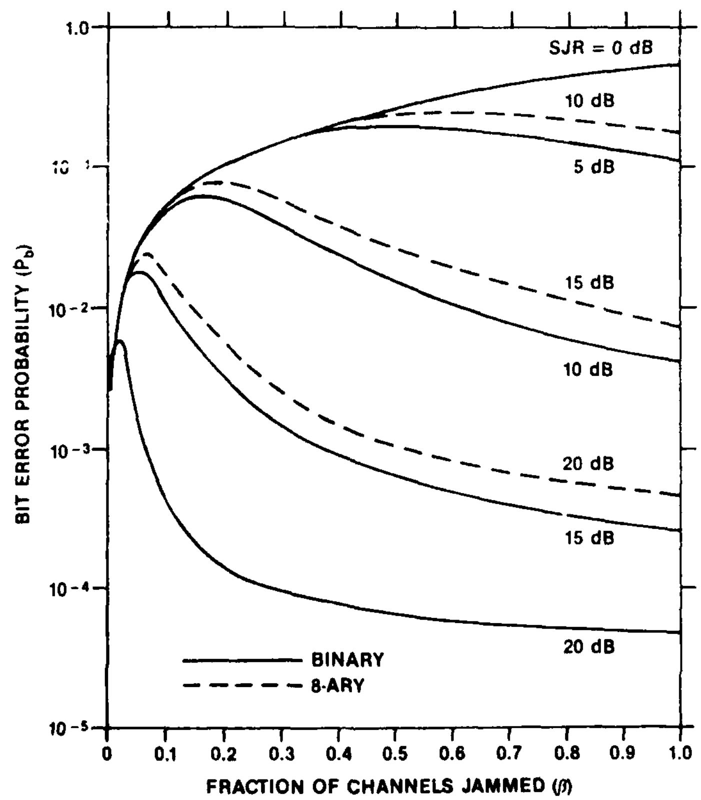


Fig.6.7 P_b versus β for SNR = 13.35 dB under multitone noise jamming. P_b as a function of $\beta = 8\gamma$ for 8-ary NCFSK in the presence of Case 3 MT jamming without, and with, system noise where SNR = 13.35 dB. Break point is at 8/SJR.

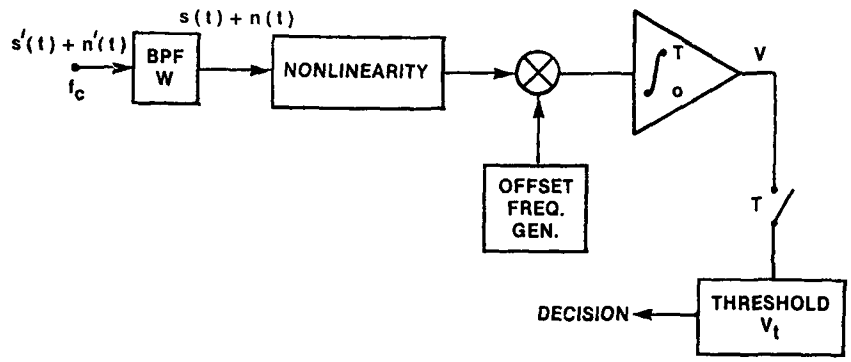


Fig.6.8 Basic form of a general detector for SS signals

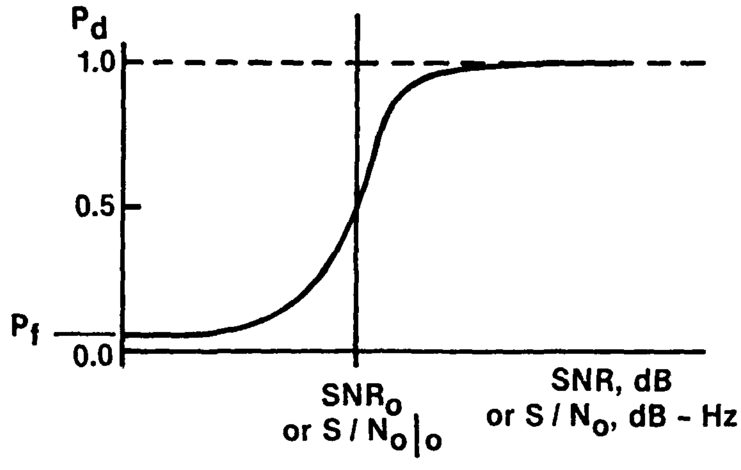


Fig.6.9 A typical "S" curve for the general detector of Fig.6.8 P_f fixed.

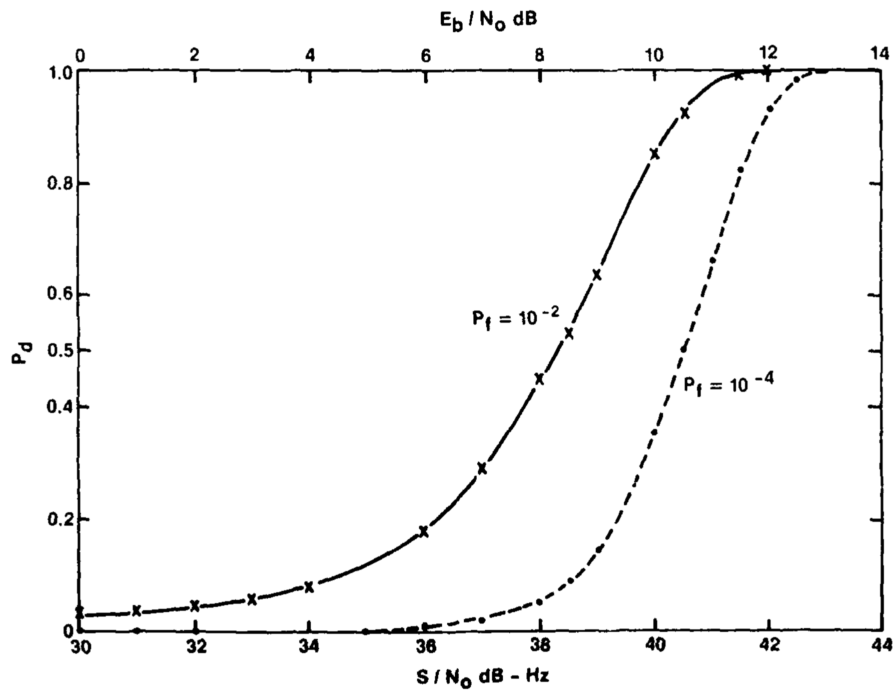


Fig.6.10 Values of P_d versus S/N_o E_b/N_o for c = 1, W = 10MHz, T = 1s, and 2 values of P_f.

Table 6.2
Typical values of W for various combinations of SS waveform and detector types

SS WAVEFORM	DETECTOR TYPE	W	c
FH	Wideband Radiometer[6.7]	W_{SS}	1.0
	Delay-and-Multiply Hop-Rate-Detector		
FH	Filter Bank Combiner [6.10], [6.11]	$\approx R_h$	1.00
DS	Wideband Radiometer,	$\approx 2R_C$	1.1
	Squaring Carrier Detector		1.0
	Relay-and-Multiply Chip-Rate Detector [6.9]		2.5
DS	Squaring Chip-Rate Detector	$< 2R_C$	
DS/FH Hybrid	Wideband Radiometer	W_{SS}	1.0

Table 6.3
Comparison of low bit-rate speech coding schemes

Coder Type	Bit rate kb/s	Complexity MIPS	Delay ms	Quality	MOS
Pulse-code modulation	64	0.01	0	High	>4
Adaptive differential pulse-code modulation	32	0.1	0	High	
Adaptive subband coding	16	1	25	High	
Multipulse linear predictive coding	8	10	35	Communication	>2
Stochastically excited linear predictive coding	4	100	35	Communication	
LPC vocoder	2	1	35	Synthetic	<2

APPENDIX 6A
A SUGGESTED APPROACH FOR ESTIMATING AND COMPARING THE ANTI-JAM CAPABILITIES OF SATELLITE COMMUNICATIONS LINKS

This Appendix suggests a simplified format for estimating and comparing the potential anti-jam capabilities of on-board signal processing satellite and hard-limited transponder satellite communications links. It does not require knowledge of the detailed engineering specifications. It is based entirely on fundamental parameters such as data rate, spread bandwidth, etc. The results of the comparisons are stated in terms of the ratio of jammer size to terminal size.

1. Signal Processing Satellites

The analytical model of a spread-spectrum AJ communications link through a singal processing satellite is shown in Fig.1. The uplink data are received, demodulated to digital bits and then retransmitted on the downlink --usually in a totally different signal format. The uplink and downlink are therefore independent. To a good approximation, the end-to-end bit error rate is the sum of the uplink and downlink bit error rates.

We will analyze only the uplink; the downlink analysis is identical (although the parameters are usually quite different). The uplink is usually the more critical of the two because it can be attacked by earth-based jammers located anywhere within sight of the satellite, whereas the downlink receiver is susceptible only to nearby jammers.

Referring to Fig. 1, the signal and jammer powers at node "u" are S_u and J_u, respectively. The receiver noise power (assumed white) is $N_u \cdot W$ where $N_u = kT_{sys}$. T_{sys} is the system noise temperature and W is the spread bandwidth of the link. The carrier to interference-plus-noise ratio at node "u" is

$$CNR_u = \frac{S_u}{J_u + N_u \cdot W} \qquad (1\text{-}1)$$

For a first order analysis we assume that the jammer has a white noise spectrum, i.e., its power is uniformly distributed over W with density J_u/W. It is generally accepted (see for example, Ref.1) that a properly designed AJ link under sophisticated jamming attack (other than white noise) can perform within a few dB of the white noise case. This is accomplished by using coding, interleaving and modulation techniques which together achieve a high degree of time and frequency diversity. For this reason the white noise performance is used as a general measure of the overall robustness of an AJ link.

The white noise performance is calculated as follows: for any communications link, the criterion for successful communications is that,

$$(E_b/N_o) \geq (E_b/N_o)_r \qquad (1\text{-}2)$$

where E_b is the received energy per information bit and N_o is the spectral density of the noise-plus-interference in watts per Hz.
$(E_b/N_o)_r$ is the ratio required for communications at the specified bit error rate(such as 10^{-3} for example). Typical values for $(E_b/N_o)_r$ range from about 4 to about 40. A value of 10 (10 dB) is representative. If R is the data rate then, for the uplink,

$$E_b = S_u/R \qquad (1\text{-}3)$$

The noise-plus-interference power density is

$$N_o = N_u + J_u / W \qquad (1\text{-}4)$$

Therefore the criterion for succesful communications (Eq.1) can be rewritten as

$$\frac{S_u}{J_u + N_u \cdot W} \geq \frac{R(E_b / N_o)_r}{W} \qquad (1\text{-}5)$$

From this we define the "Ideal Jammer Standoff Ratio" as

$$JSR_o = \frac{W}{R(E_b / N_o)_r} \qquad (1\text{-}6)$$

If we ignore the thermal noise term ($N_u \cdot W$) in Eq. 1-5, which is a very good approximation in a jamming scenario, then the criterion for successful communication is

$$\frac{J_u}{S_u} \leq JSR_o \qquad (1\text{-}7)$$

Example

Typical values for an uplink might be
1) Spread Bandwidth W = 1 GHz
2) Data Rate R = 10^3 bps
3) Required (E_b/N_o) $(E_b/N_o)_r$ = 10 (10 dB)

in which case

$$JSR_o = \frac{10^9}{10^3 \cdot 10} \quad 10^5 \text{ or } 50 \text{ dB} \qquad (1\text{-}8)$$

The physical interpretation of this ratio is that, at the uplink receiver input, the jammer must be 50 dB more powerful than the signal in order to be capable of disrupting communications.

The above statement holds true for jammer and signal powers as measured at the uplink receiver. In order to relate these quantities to the actual sizes of the jammer and user terminal, we must take into account the effects of antenna discrimination by modifying Eq. 1-7 as follows:

$$\frac{ERP_J}{ERP_s} \leq JSR_o \left(\frac{G_S}{G_J} \right) \qquad (1\text{-}9)$$

Here ERP means effective radiated power and is an indication of the physical size of the jammer or terminal. G_J is the uplink receive antenna gain in the direction of the jammer and G_S is the gain in the direction of the signal source. If the uplinks has an earth coverage antenna no advantage is gained from antenna discrimination. Of course one strives in designing the system to have low gain on the jammer and high gain on the user

Two techniques for achieving significantly greater robustness through antenna discrimination are:

1) Spot beam antennas pointed at the friendly terminal, for which the jammers are hopefully in the antenna sidelobe regions and therefore suppressed by 20 to 40 dB and

2) A nulling antenna such as implemented on DSCS-III or Skynet 4. In the latter case, the nulling antenna must be designed so that it does not null the signal as well as the jammer.

In either case, it is reasonable to expect that one might achieve 30 dB of additional robustness through antenna discrimination. In that case the previous example would work out to:

$$\frac{ERP_J}{ERP_S} = 50\text{ dB} + 30\text{ dB} = 80\text{ dB}$$

or, in other words, the jammer would have to have an effective radiated power 80 dB greater than the terminal in order to disrupt communications.

The ERP of a terminal or jammer is the product of the radiated RF power and the transmit antenna gain, which is closely related to the physical size of the antenna. An 80 dB difference in ERP represents a very considerable difference in physical size.

This completes the description of the performance of the uplink with white noise jamming. Notice that the formulation depends only on fundamental parameters such as data rate, spread bandwidth, $(E_b/N_o)_r$ and antenna gain. It is independent of the detailed design of signal waveforms etc. The downlink can be described in exactly the same way as the uplink.

2. Hard-Limiting Transponder Satellite

The analytical model for an AJ communications link through a hard-limiting transponder satellite is shown in Fig. 2. The transponder is modelled as a hard limiter, a frequency translator, a band pass filter of bandwidth W, and a power amplifier with output power P_t, which is constant because of the hard limiter. At the uplink receiver (node "u"), S_u, J_u and N_uW are the signal, jammer and thermal noise powers respectively. The TWTA output power, P_t, consists of signal power plus retransmitted noise and jammer powers. Let x be the fraction of signal power contained in P_t. The retransmitted noise-plus-jammer power is therefore $(1-x)P_t$. The carrier to noise-plus-interference ratio (CNR) at node "t" is therefore

$$CNR_t = \frac{x}{1-x} \qquad (2\text{-}1)$$

Hard limiting transponders are characterized by a suppression ratio, Γ, which is defined in terms of the input and output CNR's(Refs. 2 and 3).

$$\Gamma = \frac{CNR_t}{CNR_u} \qquad (2\text{-}2)$$

Γ itself is a non-linear function of CNR_u and depends in a complicated way on the spectra of S_u, J_u and N_u. The relationships between x and CNR_u are:

$$\Gamma \cdot CNR_u = \frac{x}{1-x} \qquad (2\text{-}3)$$

and

$$x = \frac{\Gamma\, CNR_u}{1 + \Gamma\, CNR_u} \qquad (2\text{-}4)$$

Let L be the net loss between nodes "t" and "d" on the downlink. (L includes the antenna gains at both ends of the link and the path loss.) From Fig. 2, it is clear that the carrier to interference-plus-noise ratio at the downlink receiver node "d" is

$$CNR_d = \frac{xLP_t}{(1-x)LP_t + N_dW + J_d} \qquad (2\text{-}5)$$

where $(1-x)LP_t$ is the retransmitted noise-plus-interference.

As before with the signal processing satellite link, the criterion for successful communications is that, at the downlink receiver (node d)

$$(E_b/N_o)_d \geq (E_b/N_o)_r \qquad (2\text{-}6)$$

The relationship between (E_b/N_o) and CNR at node d is

$$(E_b/N_o)_d = CNR_d\,(W/R) \qquad (2\text{-}7)$$

where W is the spread bandwidth and R is the data rate. The criterion in Eq. 2-6 can be written (from Eqs. 2-5 and 2-7)as

$$\frac{xLP_t}{(1-x)LP_t + N_dW + J_d} \geq (E_b/N_o)_r(R/W) \qquad (2\text{-}8)$$

It is convenient to rewrite this criterion in terms of two normalized variables defined as follows:

1. The Ideal Jammer Standoff Ratio

$$JSR_o = \frac{W}{R(E_b/N_o)_r} \qquad (2\text{-}9)$$

This ratio is the measure of performance of a spread spectrum communications link with white noise jamming, which was introduced in the analysis of the singal processing satellite (Eq. 1-6).

2. The Downlink Margin, (defined as if there were no uplink jamming)

$$M = \frac{LP_t}{(N_dW + J_d)} \cdot \frac{W}{R(E_b/N_o)_r} \qquad (2\text{-}10)$$

With these definitions along with Eq. 2-4, we can define an actual jammer standoff ratio, JSR_u, for the uplink terms of JSR_o and M as follows.

$$JSR_u = (CNR_u)^{-1} \qquad (2\text{-}11)$$

and

$$JSR_u = \frac{\Gamma \cdot JSR_o \cdot (M-1)}{JSR_o + M} \qquad (2\text{-}12)$$

This equation results from straightforward algebraic manipulation of Eq. 2-8. The derivation is given in the Annex.

So far we have assumed implicitly that the uplink jamming is white noise and is retransmitted on the downlink. If the uplink jammer is a tone signal, then the only retransmitted interference occurs exactly at the translated tone frequency and its harmonics. Elsewhere in the band the equivalent experession to Eq. 2-12 is

$$JSR_u = \Gamma \cdot (M-1) \qquad (2\text{-}13)$$

Equations 2-12 and 2-13 can be interpreted with the help of the curves in Fig. 3. In the case of white noise jamming (Eq. 2-12) there are two regimes: $M >> JSR_o$ and $M << JSR_o$. In the first regime, the AJ performance asymtotically approaches the ideal performance (JSR_o) within a factor of Γ. In the second regime, $M << JSR_o$, the performance is determinet entirely by the available downlink margin. This is the so-called "power robbing" regime. In the tone jamming case, the AJ performance is only dependent on M. Some examples follow.

Example 1 (Power Robbing Regime)

Given:

1) Spread Bandwidth	W	= 1 GHz
2) Data Rate	R	= 10^3 bps
3) Required (E_b/N_o)	$(E_b/N_o)_r$	= 10 (10 dB)
4) Downlink Margin	M	= 10 (10 dB)

(The downlink margin, M, includes any allowance for white noise downlink jamming J_d.)

The suppression ratios for various cases are discussed in Refs. 2 and 3. For tone and white noise jamming they are as follows:

5) Tone Jamming	Γ = 1/4(-6 dB)
6) White Noise Jamming	Γ = 3/4(-1.25 dB)

Calculate:

I) Ideal Jammer Standoff Ratio

$$JSR_o = \frac{W}{R(E_b / N_o)_r} = 10^5 \quad (50\text{ dB})$$

Since $M << JSR_o$, we are in the power robbing regime. For tone or white noise uplink jamming, we have

2) $JSR_o = \Gamma.(M-1) \approx 3$ dB for tones or ≈ 8 dB for white noise.

This means that a tone jammer with a power only 3 dB above the signal power, as measured at the uplink receiver (node u), will disrupt the link. (Antenna discrimination could improve this performance as indicated previously).

Example 2 (Indermediate Case)

Given:

1) W	= 10 MHz
2) R	= 10^3 bps
3) $(E_b/N_o)_r$	= 10 (10 dB)
4) M	= 10^3 (30 dB)
5) Γ	= 1/4 for tone jamming
6) Γ	= 3/4 for white noise jamming

Calculate:

$$JSR_o = \frac{W}{R(E_b / N_o)_r} = \frac{10^7}{10^3 \cdot 10} = 10^3 \text{ (30 dB)}$$

Since $M = JSR_o$, we are at the boundary of the two regimes in Fig. 3

2) For Tone Jamming

$$JSR_u = \Gamma(M - 1) = -6\text{ dB} + 30\text{ dB} = 24\text{ dB}$$

3) For White Noise Jamming ($M >> JSR_o$)

$$JSR_u = \frac{\Gamma \cdot JSR_o (M - 1)}{JSR_o + M} = 3/8 \times 10^3 = 26\text{ dB}$$

Note that this level of AJ performance is attained by having a large downlink margin available.

Example 3

Same as example 2 except that $R = 10^4$bps. $JSR_o = 10^2$(20dB). $M = 10^3$ as before. For tone jamming JSR_u = 24dB as before and for white noise jamming

$$JSR_u = \frac{3/4 \cdot 10^2 \cdot 10^5}{10^2 + 10^5} \cdot \frac{3}{4} 10^2 = 19\text{ dB}$$

Note that the performance approach the ideal performance for a link with that particular choice of spread bandwidth and data rate.

A final observation: in order to achieve a specified JSR_u, say 50 dB, with a hard limiting transponder satellite, one must require that JSR_o > 50dB and M > 50dB. A signal processing satellite only requires the first of these conditions, i.e., JSR_o > 50dB for the uplink. This is the fundamental advantage of signal processing satellites. One does not need to squander precious downlink transmitter power in the form of excess margin, M in order to achieve substantial AJ protection. Of course the disadvantage of signal processing lies in the increased complexity of the satellite electronic required.

REFERENCES

[1] Simon, M.K., Omura, J.K., et al., "Spread Spectrum Communications", Vol.I, Computer Science Press, Inc., 1985, Chapter 5.

[2] Gagliardi, R.M. "Satellite Communicaitons", Van Nostrand Reinhold Co., 1984, pp.166 and 202.

[3] Spilker, J.J., "Digital Communications by Satellite", Prentice Hall, Inc., 1977, pp. 226-230.

ANNEX to Appendix 6A

Derivation of Eq. 2-12 and Eq. 2-13

We first want to express the criterion for successful communications, Eq. 2-6, in terms of the uplink carrier-to-noise ratio, CNR_u.

Start with the inverse of Eq. 2-8 and substitute Eq. 2-9 for the right-hand site and Eq. 2-4 for x.

$$\frac{1}{\Gamma CNR_u} + \frac{(NdW + J_d)(1 + \Gamma CNR_u)}{LP_t \Gamma CNR_u} \le JSR_o \qquad \text{(A-1)}$$

Multiply by ΓCNR_u and divide by JSR_o. Substitute Eq. 2-10 for M.

$$\frac{1}{JSR_o} + \frac{1}{M} (1 + \Gamma \cdot CNR_u) \le \Gamma CNR_u \qquad \text{(A-2)}$$

Solve for CNR_u.

$$CNR_u \ge \frac{M + JSR_o}{\Gamma \cdot JSR_o \cdot (M - 1)} \qquad \text{(A-3)}$$

This is an alternative form of the criterion for successful communications. We define the jammer standoff ratio for the system as the maximum jammer-plus-noise to signal ratio that the system can tolerate and still communicate successfully, i.e.,

$$JSR_u = \text{Max}\left[(CNR_u)^{-1} \right] \qquad \text{(A-4)}$$

or

$$JSR_u = \frac{\Gamma\, JSR_o\, (M - 1)}{M + JSR_o} \qquad \text{(A-5)}$$

This expression pertains to white noise jamming. For tone jamming, we modify Eq. 2-8 as follows.

$$\frac{xLP_t}{N_d W + J_d} \ge (E_b/N_o)_r \quad (R/W) \qquad \text{(A-6)}$$

This expression is valid everywhere in the band except exactly at the translated tone frequency or its harmonics. Eq. 2-13 then follows straightforwardly.

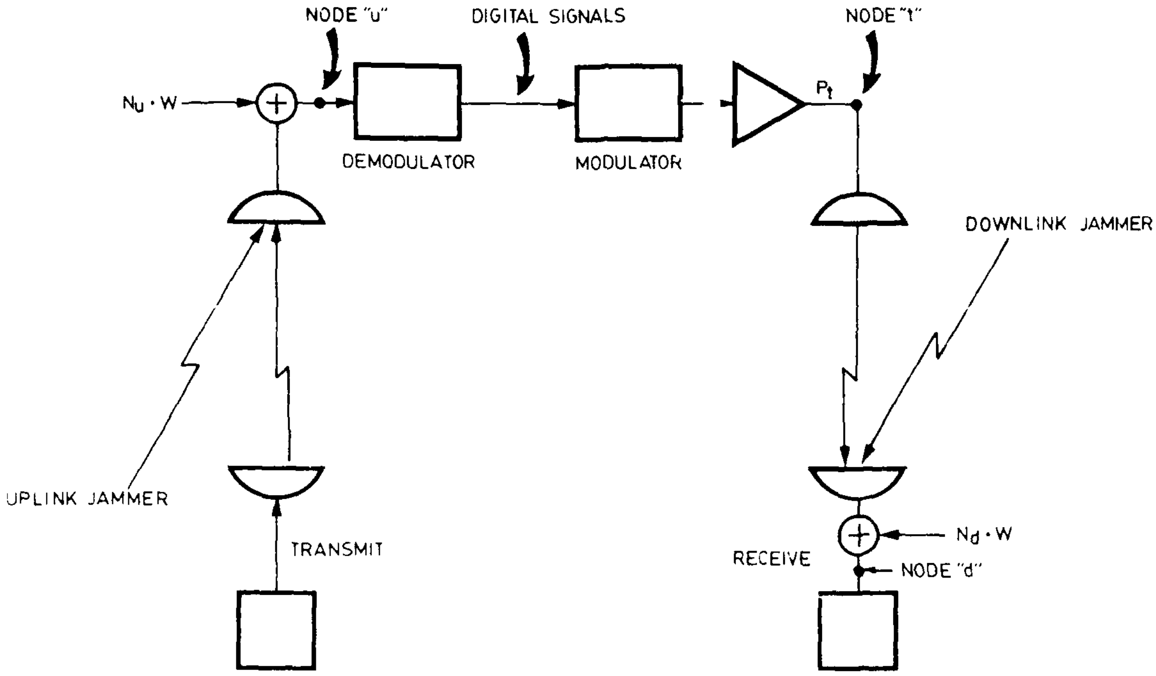


Fig 1 Analytical model of signal processing satellite communications link

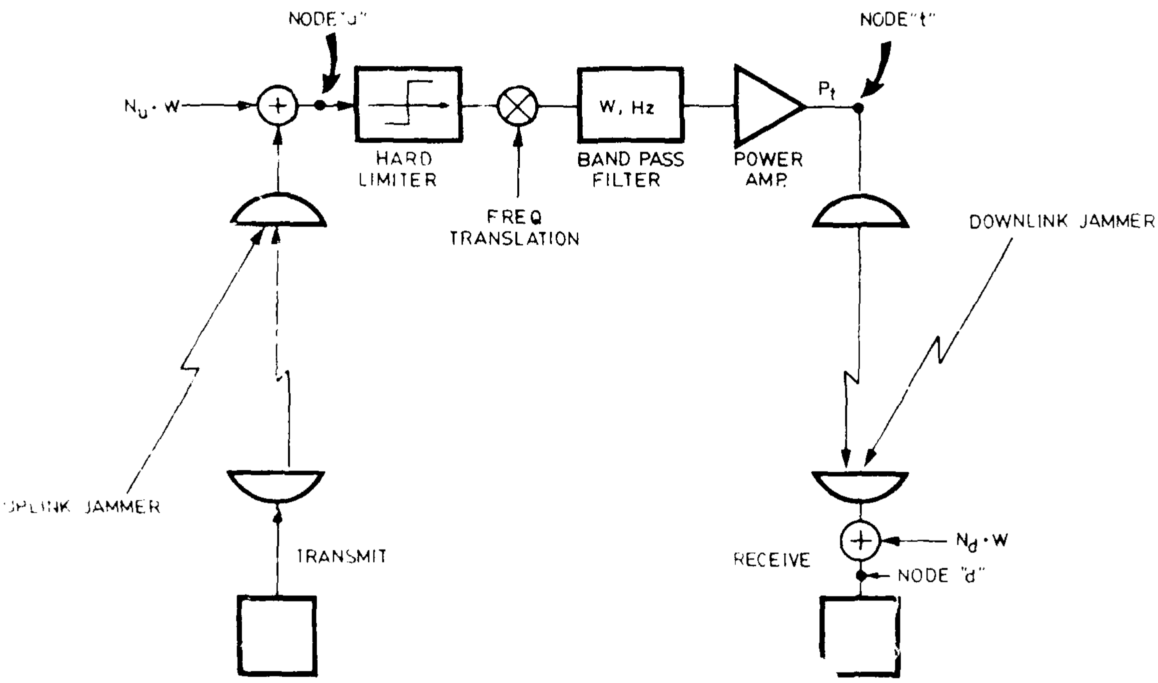


Fig 2 Analytical model of hard limiting

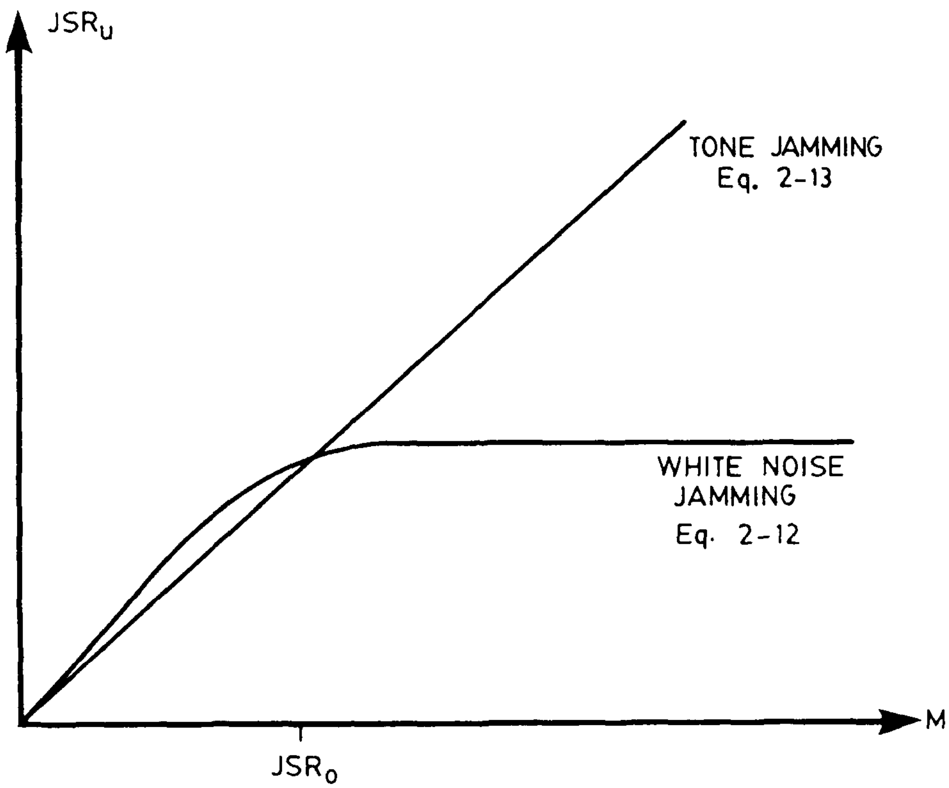


Fig 3 Jammer standoff ratio, JSR , for a hard-limited transponder as a function of downlink margin, M

APPENDIX 6B
FEC CODING FOR SATCOM

To demonstrate the efficiency of FEC techniques for different applications of interest for WG 13 some examples from the wide range of well-known schemes, are selected.

Considered are:

a) Communication in an additive white Gaussian noise channel (AWGNC)
b) Communication in a Rayleighfading channel (mobile satellite communications)
c) Direct Sequence Spread-Spectrum communication in a channel with jamming
d) Frequency hopping Spread-Spectrum communication in a channel with jamming

Areas of current investigations:
- Coded Modulation (Trellis coded Modulation, Block coded Modulation) for fixed as well as mobile satellite communications.
- 'channel adaptive' coding (variable rate coding with punctured codes)

1. FEC For An AWGNC

The required signal-to-noise ratios E_b/N_o to achieve a bit-error-rate $P_b = 10^{-5}$ for several coding schemes with binary PSK or QPSK are listed in Table 1. Table 2. shows some results for noncoherently detected M-ary FSK.

For coded 8-PSK with Ungerboeck convolution coding constraint length 3, $R = 2/3$, soft decision Viterbi decoding, $E_b/N_o \approx 7.1$ dB is required for $P_b = 10^{-5}$. Compared with uncoded QPSK this corresponds to a coding gain of 2.5 dB, whereas the bandwidth of both systems is the same.

E_b : energy per information bit
N_o : one-sided noise power density
R : coding rate in inf-bit per coded bit

2. Coding for a Rayleighfading channel

In some applications (mobile SATCOM) signal fading can severely degrade the error rate performance of a system. Considering a frequency-nonselective slowly Rayleighfading channel the bit-error-rates for uncoded binary PSK is

$$P_b = \frac{1}{2}\left[1 - \sqrt{\frac{E_b/N_o}{1 + E_b/N_o}}\right]$$

or for uncoded binary FSK

$$P_b = \frac{1}{2 + E_b/N_o}$$

with

E_b: average received energy-per-bit

2.1 Diversity, Coding

The performance of a system in a fading channel can be greatly improved by providing some type of diversity; that is, by providing m independent transmissions for each transmitted symbol. The diversity technique may be regarded as a simple repetition code of rate $1/m$.

Using interleaving/deinterleaving the received channel symbols can be treated as independent random variables.

In Table 3 the required signal-to-noise ratios achieve $P_b = 10^{-5}$ are summarized for m-diversity binary PSK (maximal-ratio combining) and m-diversity noncoherently detected M-ary FSK (square-law combing).

Applying more sophisticated FEC techniques, interleaving/deinterleaving is also assumed.

If information about the channel-state is available at the receiver, this information can be utilized for decoding.

The results for several coded schemes with PSK and noncoherently detected FSK on a Rayleigh fading channel are given in Table 4.

3. Coding for Frequency-hopped Systems

3.1 Parameters

S : received signal power

R_b : information bit rate

E_b : energy per information bit

$E_b = S/R_b$

R : code rate, in information bits/coded (transmitted) bit

r : overall coding rate in information bit/M-ary channel
$r = R_b/R_c$

R_s : M-ary symbol rate ($M = 2^K$)
$R_s = R_b/(K.R)$

R_h : frequency hop rate

R_c : chip rate, $R_c = \max(R_H, R_s)$

W_s : system bandwidth
$W_s = N_t . R_c$ (orthogonality)

N_t : total number of frequencies

P_G : Processing gain:
$P_G = W_s/R_b$

Slow frequency-hopping (SFH):

$$R_s \geq R_h$$

including symbol-interleaving (a symbol is distributed over m hops) the chip rate is

$$R_c = m.R_s \geq R_h$$

If $R_s = l.R_h$, then the energy in one hop is $E_h = E_b.l.K.R$

Fast frequency-hopping (FFH):

$$R_h > R_s$$

with m hops per symbol

$$R_h = m.R_s$$

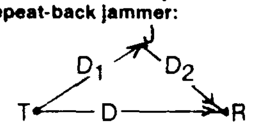
and the energy per hop

$E_h = E_b . K . R / m = E_c$

- The performance achieved with interleaved SFH is the same as with FFH.
- Since coherent modulation like Mary-PSK, MSK, ... as well as differentially coherent modulation DPSK, DMSK require 'trackable' phase or even almost constant phase over at least two symbols, these types of modulation might be only suitable for SFH-systems. In the following noncoherent M-ary FSK is considered.

3.2 Jammers for FH-Systems

3.2.1 Repeat-back jammer:



The time duration T_J in which the jammer must react is

$$T_J < \frac{1}{C}\left[D_1 + D_2 - D\right] + T_H$$

3.2.2 Broadband noise jammer (BBNJ):

The jammer transmits broadband noise over the total Spread-Spectrum band with a constant power J. This is equivalent to an AWGNC, with an effective noise power spectral density

$$N_J = J / W_s$$

Therefore the signal-to-jammer noise ratio is

$$\frac{E_b}{N_J} = (S / J) \bullet (W_s / R_b) = (S / J) \bullet P_G$$

3.2.3 Partial-band noise jammer (PBNJ):

The jammer restricts its total power J to a fraction $\rho (0 < \rho \le 1)$ of the spread-spectrum band W_s. The free parameter ρ is optimized maximizing *the bit-error-rate* P_b*(BER) of the FH system. The* jammed band is assumed to be hopped randomly with a rate R_h but fast enough to avoid anti-jam reaction by spread-spectrum system

3.2.4 Band multitone jammer (BMTJ):

The jammer devides its total power J into Q distinct random phase CW tones of equal power. These are distributed over W_s. Analysis show, that the best multitone strategy is, to leave at least M-1 unjammed FH-slots between each of its Q available tones, so that no M-ary subband can contain more than $n = 1$ jamming tone. As for the partial-band noise jammer it is assumed, that the tones are varied randomly, so that the spread-spectrum system cannot avoid the jammed subbands. The jammer is free to optimize $\alpha = (S/J)\,Q$ within $0 < \alpha < 1$ for maximizing the BER of the jammed system

From analysis it turns out, that the bandmultitone, $n = 1$ jammer can be regarded as the overall worst-case jammed for FH-systems

3.2.5 *Independent multitone jammer (IMTJ):*

The Q tones are distributed over the available N_t frequencies. The analysis show, that this strategy is almost as efficient as the worst-case bandmultitone $n = 1$ jammer

3.3 Performance of uncoded FH/MFSK with jamming

For the analysis thermal noise is neglected

3.3.1 FH/MFSK, no diversity

It is assumed that at least one M-ary-symbol is transmitted per hop(SFH) without symbol interleaving. The bit error rates for the different jammers are:

BBNJ:

$$P_b \le \frac{M}{4} e^{-\frac{K}{2} E_b / N_J}; \quad M = 2^K$$

PBNJ:

For the worst-case jammer (optimized ρ)

$$P_b = \begin{cases} \frac{M}{4} e^{-\frac{K}{2} E_b / N_J} & ; E_b/N_J \le \gamma \ (\rho = 1) \\ \frac{\beta}{E_b/N_J} & ; E_b/N_J \ge \gamma \left(\rho = \frac{\gamma}{E_b/N_J}\right) \end{cases}$$

with $\beta(K) \le 0.36$, $\gamma(K) \le 3dB$

BMTJ(n=1):

For worst case jammer (optimized $\alpha < 1$)

$$P_b = \begin{cases} 1/2 & ; E_b/N_J \le M/K, \left(\alpha = \frac{K}{M} E_b/N_J\right) \\ \frac{M}{2KE_b/N_J} & ; E_b/N_J \ge M/K, (\alpha = 1) \end{cases}$$

IMTJ :

$$P_b \approx \frac{M/(2K)}{E_b/N_J}$$

3.3.2 FH/MFSK with time-diversity

One of the simplest (coding) techniques to improve the performance of system is to subdivide each information symbol *into equal energy subsymbols which are transmitted over* independent channel states. In the context of FH/MFSK, an M-ary symbol is partitioned into m subsymbols 'chips' and each chip transmitted on a different hop using (FFH) or SFH with interleaving.

PBNJ:

For the worst-case jammer (optimized ρ) and for an optimized m, the bit error rate

$$P_b \approx \frac{M}{4} e^{-\beta E_b/N_J}; \quad \frac{E_b}{N_J} \ge \frac{1}{\beta}$$

$$m_{opt} = \gamma\, E_b/N_J; \quad \gamma \approx K/4$$

BMTJ:

For the worst-case jammer (optimized α) and for m optimized, *the bit error rate*

$$P_b \le \frac{M}{2} e^{-\delta E_b/N_J}; \quad \frac{E_b}{N_J} \ge \frac{1}{\delta}$$

$$m_{opt} = \delta\, E_b/N_J; \quad \delta(K) < 0.47$$

In Table 5 results for FH/MFSK with and without diversity are summarized

3.4 PERFORMANCE OF CODED FH/MFSK

3.4.1 Coding with and without diversity

Coding schemes, considered:

a) binary convolutional code, constrained length 7, rate 1/2, soft decision energy detection/decoding with perfect channel state information.
b) M-ary conv. code, constr.length 7, rate 1/2; 1/3; soft decision energy detection/decoding with perfect channel state information.
c) dual-K conv. codes, rate 1/2; 1/3; soft decision energy detection/decoding with perfect channel state information.
d) Reed-Solomon (n,k) codes over GF(2^Q), hard decision on the received symbols assumed.
e) concatenated coding with RS (n,k) outer code and dual-K inner code, soft decision energy detection/decoding with perfect channel state information for the inner code assumed.
f) concatenated coding RS(n,k) outer code and
   binary conv.code, constr. length 7, rate 1/2 (M=2)
   4-ary conv. code, constr. length 7, rate 1/2 (M=4)
   8-ary conv. code, constr. length 7, rate 1/3 (M=8)

(inner code soft decision inner decoding with perfect channel state information).

Using additional diversity, the number m of diversity channels is optimized.

The results are summarized in Table 6.

4. CODING FOR DS SPREAD-SPECTRUM SYSTEM

Parameters

R_b : information bit rate
E_b : energy per information bit
R : coding rate
R_c : chip rate
E_c : energy per chip
W_s : system bandwidth
P_G : processing gain

$$P_G = \frac{W_s}{R_b}$$

4.1 Jammers for DS-Systems

Broad-band noise jammer (BBNJ):
The jammer transmits a signal $J(t)$ with constant power J. The spread-spectrum system can be characterized by its signal-to-jamming noise ratio

$$\frac{E_b}{N_J} = \left(\frac{S}{J}\right)\left(\frac{W_s}{R_b}\right) = \left(\frac{S}{J}\right) P_G$$

Pulse jammer (PJ):

The jammer transmits a signal, but for only a fraction $\rho \cdot R_b^{-1}$ of time, with the peak power Işlem J/ρ. J denotes the time averaged jammer power

4.2 Performance of DS-Spread Spectrum with jamming

In the following it is assumed, that the thermal noise can be ignored.

4.2.1 Performance of uncoded DS/PSK

BBNJ:

For broadband noise jamming and binary PSK or QPSK the bit error rate

$$P_b = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{E_b/J}\right).$$

For large P_G the above bit error rate also applies to most constant power jammer waveforms (for performance, see AWGNC).

PJ:

Since the jammer is free to optimize ρ, maximizing the bit error rate of the spread spectrum system, for a DS System with binary PSK or QPSK:

$$P_b = \begin{cases} \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{E_b/N_J}\right) & ; E_b/N_J \le 0.71 \\ \frac{0.83}{E_b/N_J} & ; E_b/N_J > 0.71 \end{cases}$$

$$P_{WC} = \begin{cases} \frac{0.71}{E_b/N_J} & ; E_b/N_J > 0.71 \\ 1 & ; E_b/N_J \le 0.71 \end{cases}$$

The required E_b/N_J to achieve $P_b = 10^{-5}$

$E_b/N_J = 39.2$dB

4.2.2 Performance of coded DS/PSK

BBNJ:
For constant power jammers the results for the AWGNC can be applied.

PJ:
Using a conv. code constr. length 7, rate 1/2 with soft decision Viterbi decoding with no channel state information or with perfect channel state information, the E_b/N_J required for $P_b = 10^{-5}$ is

E_b/N_J = 10dB (no channel state)
E_b/N_J = 6.5 dB (with channel state)

5.0 SUMMARY

Summary of the results presented in Sections 1-4 above is given in Table 7.

Table 1
Different coding schemes with binary PSK or QPSK in an AWGNC

Type of code	required E_b / N_o *) to achieve $P_b = 10^{-5}$ [dB]	coding gain *) [dB]
uncoded	9.6	– – –
conv. code constr. length 7, soft decision, Viterbi decoding		
R = 1/2	4.2	5.4
R = 2/3	4.6	5.0
R = 3/4	5.0	4.6
punctured:		
R = 2/3	5.0	4.6
R = 3/4	5.6	4.0
conv. code, constr. length 3 soft decision Viterbi decoding		
R = 1/2	5.8	3.8
R = 2/3	6.5	3.0
R = 3/4	6.7	2.9
extend. Golay (24,12), hard decision (soft decision)	7.5 (5.6)	2.1 (4.0)
BCH – code, hard decision decoded		
(127,64), R≈ 1/2	6.2	3.4
(511, 259), R≈ 1/2	5.5	4.1
RS – code, hard decision decoded		
(31,17), R≈ 0.55	6.7	2.9
(31,25), R≈ 0.8	7.2	2.4
(255,192), R = 3/4	6.0	3.6
concaten. code RS (255,223) outer code conv. code constr. length 7, Rate1/2 (inner code) R = 0.44	2.6	7.0

*) the given values are approximations

Table 2
Different coding schemes with noncoherent M-ary FSK in an AWGN

Type of code	required E_b/N_o for $P_b = 10^{-5}$ in dB					Coding gain in dB
	M=2	M=4	M=8	M=16	M=32	
uncoded	13.4	10.6	9.1	8.1	7.4	---
binary conv. code constr. length 7, rate 1/2	9.7	---	---	---	---	3.7
M-arry conv. code constr. length 7, rate 1/2	---	7.5	---	---	---	3.1
8-arry conv. code constr. length 7, rate 1/3	---	---	7.1	---	---	2.0
RS (31 , 15)	---	---	---	---	5.9	1.5
RS (31 , 23)	---	---	---	---	5.5	1.9

Table 3
Delivery performance in a frequency — nonselective, slowly Rayleighfading channel

	required $\overline{E_b}/N_o$ for $P_b = 10^{-5}$ in dB for m-Diversity			
Modulation	m=1	m=2	m=4	m=8
BPSK	44	24.4	16.4	14.
2-FSK *)	50.	31.	22.	19.
4-FSK *)	≤ 48.	30.	22.	---
8-FSK *)	≤ 48.	28.	19.	---
16-FSK *)	≤ 48.	27.	18.	---
32-FSK *)	≤ 48.	27.	17.	---

*) noncoherent demodulation

Table 4
FEC performance in a frequency — nonselective, slowly Rayleighfading channel

Type of Code	Modulation technique	$\overline{E}_b/N_o$ in dB required for $P_b = 10^{-5}$	Coding gain in dB
conv. Code R = 1/2 constr. length 3, soft decision Viterbi decoding	2-FSK *)	17.5	32.5
conv. Code R = 1/2 constr. length 5, soft decision Viterbi decoding		15.5	34.5
same code combined with m = 4 diversity R = 1/8		13.0	37.0
dual - k conv. code R = 1/2	8-FSK *)	16.5	31.5
conv. code constr. 7 R = 1/2 soft decision Viterbi decoding no channel state (one bit channel state)	BPSK	14.0 (11.0)	30.0 (33.0)
conv. code constr. length 3 R = 1/2, soft decision Viterbi decoding no channel state (one bit channel state)		19.0 (14.0)	25.0 (30.0)
RS (15.7) hard decision decoded		16.0	28.0

*) noncoherent demodulation

Table 5
E_b/N_J in dB required to achieve $P_b = 10^{-5}$ for FH/MFSK with and without diversity and for different types of jammer

Type of Jammer	M-ary FSK									
	M = 2		M = 4		M = 8		M = 16		M = 32	
	E_b/N_J	m_{opt}	E_b/N_J	m_{opt}	E_b/N_J	m_{opt}	E_b/N_J	m_{opt}	E_b/N_J	m_{opt}
BBNJ	13.4	- - -	10.6	- - -	9.1	- - -	8.1	- - -	7.4	- - -
PBNJ	45.7 16.4	- - - 11	43.7 13.6	- - - 11	42.9 12.1	- - - 12	42.6 11.1	- - 13	42.5 10.4	- - - 14
BMTJ	50.0 15.0	- - - 12	50.0 14.2	- - - 12	51.2 14.6	- - - 13	53.0 15.5	- - - 14	55.1 16.8	- - - 14
IMTJ	50.0 15.0	- - - 9	50.0 13.5	- - - 16	51.2 13.5	- - - 16	53.0 14.1	- - - 22	55.1 15.1	- - - 26

Table 6
Required E_b/N_J to achieve $P_b = 10^{-5}$ for different coding schemes with FH/MFSK signals in worst-case partial band noise and band multitone jamming

Type of code	Type of jammer	M–ary FSK M=2 E_b/N_J, dB	M=2 m_{opt}	M=4 E_b/N_J, dB	M=4 m_{opt}	M=8 E_b/N_J, dB	M=8 m_{opt}	M=16 E_b/N_J, dB	M=16 m_{opt}	M=32 E_b/N_J, dB	M=32 m_{opt}
binary conv. code	PBNJ	11.1	- -								
7, rate 1/2		10.7	2								
	BMTJ	9.7	- -								
		9.3	2								
M-ary conv. code	PBNJ			10.1	- -	9.0	- -				
7, rate 1/3, rate 1/2				8.9	2	8.3	2				
	BMTJ			10.8	- -	11.6	- -				
				9.4	2	10.7	2				
dual-K conv. code	PBNJ	16.7	-	14.4	- -	13.4	- -	12.0		12.6	-
rate 1/2		13.5	3	10.7	3	9.2	3	8.2	3	7.4	4
	BMTJ	15.8	- -	15.5	- -	16.3	- -	17.8	- -	19.6	- -
		12.0	3	11.3	3	11.6	3	12.5	3	13.8	3
dual-K conv. code,	PBNJ	14.4	- -	11.9	- -	10.6	- -	9.9	- -	9.4	- -
rate 1/3		13.4	2	10.6	2	9.2	2	8.1	2	7.4	2
	BMTJ	13.2	- -	12.7	- -	13.3	- -	14.5	- -	16.0	- -
		12.0	2	11.2	2	11.6	2	12.5	2	13.8	2
RS (255, 191)	PBNJ	17.7	- -	14.5	- -			14.5	- -		
code		13.4	4	10.8	5			8.6	5		
	BMTJ	22.0	- -	20.8	- -			20.5	- -		
		12.4	5	11.8	5			13.3	6		
concatenated	PBNJ	11.7	- -	9.4	- -			8.0	- -		
RS (255 , 191)		11.7	1	9.0	2			7.0	2		
binary conv. code,											
7, rate 1/2											
	BMTJ	10.6	- -	10.2	- -			12.5	- -		
		10.3	2	9.6	2			11.3	2		
concatenated	PBNJ	12.1	- -	9.6	- -				7.8	- -	
RS (1023 , 959)		11.4	2	8.7	2				5.9	2	
dual-K conv. code											
rate 1/2											
	BMTJ	11.0	- -	10.5	- -				14.4	- -	
		10.1	2	9.4	2				12.2	2	
concatenated	PBNJ	10.2	- -								
RS (255 , 191)		10.2	1								
binary conv. code,											
7, rate 1/2											
	BMTJ	8.7	- -								
		8.7	1								
concatenated	PBNJ			7.9	- -	6.7	- -				
RS (255 , 191)				7.9	1	6.7	1				
M-ary conv. code,											
rate 1/2, 1/3											
	BMTJ			8.6	- -	9.4	- -				
				8.6	1	9.4	1				

Table 7
Summary table of performance

Performance	E_b/N_o required for BER of 10^{-5}		
Section No:	Best uncoded /without diversity and/or worst jamming (dB)	best coded/ with diversity	
		(dB)	code
1	9.6	2.6	concaten. code
2	44	13	m = 4 diversity and *conv.* code R = 1/8
3	40	10	conv. constraint by 7 R = 1/2, soft decision
		6.5	" " with channel state info
4	42	5.9	concaten. dual - K conv. code R = 1/8

CHAPTER 7
ENVIRONMENT

7.1 PROPAGATION FACTORS

7.1.1 Introduction

Propagation for the earth-space path has been dealt with at great length and depth. Only a few broad outline generalities will be dealt with here. A brief overview is found in Spilker's book [7.1]. An in-depth Handbook [7.2] has been prepared for NASA on propagation for frequencies below 10 GHz. A number of the papers in the AGARD Conference Preceedings [7.3] are of direct relevance to milsatcom. A NATO handbook [7.4] on millimeter wave propagation in general includes consideration of the earth-space path. An AGARD lecture series document [7.5] includes a detailed section on SATCOM propagation but with the emphasis on the lower frequency bands. A good overview is found as part of an AGARD course [7.6]. Despite its title containing "UHF/SHF", it also covers EHF propagation. It also is one of the very few unclassified writings on the effects of nuclear scintillation.

One way of viewing propagation factors is to consider three aspects: 1) the propagation phenomenon, 2) its effect on the electro-magnetic wave, and 3) the consequence to the communications link. Phenomena affecting propagation include ionization, precipitation, particulate matter, hydrometeors, and reflections. The primary effects on electro-magnetic propagation are absorption, scattering, multipath, scintillation, and depolarization. The major consequences on the received signal are attenuation and fading. Attenuation results in a decreased SNR and a corresponding decrease in performance. Fading results in a time varying SNR wherein the SNR can occasionally go a few dB above the non-faded value and will more often have decreases in SNR, well below the non-faded value.

All of the propagation effects degrading milsatcom performance tend to be highly frequency dependent. Generally speaking, the frequencies in the SATCOM allocations at 20 GHz and above suffer the most from rain attenuation but are much less susceptible to the other effects. Becauese of the current and future interest in EHF, it was decided to consider the effects of precipitation at frequencies >20 GHz separately in section 7.2, and all the other effects in this section. Only the major effects and consequences will be considered.

7.1.2 Ionospheric Scintillation

Ionospheric scintillation arises from the night-time irregularities in the F layer of the ionosphere. These irregularities are in the form of varying electron density that causes a spatially and temporally varying refractive index. This varying refractive index result in scintillation of the wave propagating through the ionosphere. This scintillation translates into a fluctuating received signal i.e. a signal with fading.

The regions most affected by ionospheric scintillation are the subouroral-to-polar latitudes, and a belt surrounding the geomagnetic equator. The scintillation varies diurnally and seasonally. Considerable effort has gone into making predictions of such scintillation based upon statictical data bases.

The scintillation, and the resulting fading is worst at the UHF bands and falls with increasing frequency. As some indication of the frequency dependence, the following table is taken from [7.1]. It shows the measured fade depth for which the signal remains above 95 % of the time. These measurements were taken during a period of large scintillation. As can be seen, the fading at UHF is severe, whereas at 7.3 GHz, and above, the fading is almost negligible.

frequency	250 MHz	23 GHz	7.3 GHz
Attenuation,dB	22 dB	2dB	<0.5dB

Because ionospheric scintillation is limited spatially and temporally, and because there are methods to reduce the effect such as spatial diversity, it has not inhibited the popularity of UHF milsatcom. Availabilities even in the worst regions still far exceed the availabilities of other long distance communications such as HF radio.

7.1.3 Multipath Fading

Multipath fading arises from the direct wave being received along with another version or versions of the same wave that has been reflect from some object(s). There is some tendency for the higher frequencies to be less prone to degradation than lower frequencies.Nonetheless all frequencies can suffer especially at low elevation angles.

A related effect is especially important for land mobile terminals. Here, objects such as trees etc. periodically block part of the wave and cause a time varying fading that is dependent upon the velocity of the vehicle and the size of the blocking object.

7.1.4 Atmospheric Attenuation

Above 10 GHz, attenuation caused by water vapor and oxygen absorption can be significant. As shown in Fig. 7.1 the attenuation is a complex function of frequency with a small peak at 22.2 GHz that is a function of relative humidity, and a very large peak at 60GHz. These curves give the total attenuation in traversing the entire atmosphere at the zenith angle. In all the EHF bands of primary interest for earth-space links (20,30, and 44GHz), this atmospheric attenuation is small compared to rain attenuation. The large absorption at 60GHz is one reason why it is a candidate for intersatellite links.

7.1.5 Rain Attenuation

There is a brief discussion of this topic in [7.1]. In general there is relatively little rain attenuation for frequencies <10GHz. For frequencies above 20GHz, see Sec.[7.2]

7.1.6 Depolarization

Depolarization is discussed in [7.2] and [7.5]. It can be caused by rainfall and by ice particles. The depolarization effect is proportional to both the attenuation and to the differential phase shift. Therefore, the depolarizaiton caused by rainfall tends to fall with decreasing frequency in the same way that attenuation does, but the phase shift tends to increase. The overall result is that depolarization tends to fall with decreasing frequency but not very rapidly. It is concluded in [7.2] that there can be depolarization effects even at frequencies below 4 GHz.

The result of depolarization is an attenuation of the received signal that fluctuates with rain rate. The use of circular polarization at SHF and above helps offset the effects of depolarization.

7.1.7 Nuclear Scintillation

The scintillation caused by high altitude nuclear blast is of

concern to SATCOM systems since such a blast can be used specifically to incapacitate radio communications of all types. There is obviously a lack of measurements from which to predict the scintillation and its effect on SATCOM. One of the few unclassified articles can be found in Sec.5 of [7.6] and a limited amount of information can be found in [7.7] along with some error correction methods tailored to handle the fading induced by the nuclear scintillation.

In summary from [7.7], EHF is much more robust to nuclear scintillation with the fading being limited to a region of a few hundred kilometers under the blast and the worst fading being over in a few tens of minutes. At the opposite end, UHF fading would be "continent" wide and last for many hours, if not days.

7.2 PRECIPITATION EFFECTS

7.2.1 Introduction

In the previous section, it was seen that precipitation in the form of rain or snow has little effect on SATCOM links except at EHF. Therefore, this section considers the effect of precipitation at 20 GHz and above in the permitted EHF bands. The EHF bands most often considered for milsatcom are 30 and 44 GHz for the uplink and 20 GHz for the downlink. One of the reasons for going to the EHF bands is that there is negligible propagation effects caused by the ionosphere and there is much more tolerance to the effects of nuclear scintillation than at lower frequencies. Unfortunately, the rain attenuation is more severe at these frequencies and is the subject of this section. The issues to be addressed are rain-rate distribution and link availability maps.

It has been found that falling snow affects the propagation at EHF very little. Therefore, only the effects of rain are considered in this section.

7.2.2 Rain-rate Climate Regions

The global model used by the CCIR for rain rate divides the world into 5 zones. A more recent model uses 8 zones and is often identified with Crane [7.8]-[7.9]. This map is show in Fig 7.2 with an expanded view of continental U.S. and southern Canada show in Fig 7.3. The rain-rate distribution for each of the 8 regions is shown in Fig 7.4. Notice that in the far north (tundra) and the nearer north (taiga), the rain precipitation is low and is in fact lower than the areas designated "arid". Therefore, these regions should be ideal for EHF operation.

Where long-term rain data are available, more refined maps are possible. Crane [7.9] gives one such map for CONUS based upon records from many stations. Similarly, the expected fraction of a year that a given rain-rate will be exceeded is given in [7.10] for 47 locations in Canada.

7.2.3 Availability

Workers from Lincoln Laboratories have combined the world rain-rate distributions with computations of attenuation arising from a given rain-rate for an earth-satellite path. The technique is described in [7.11] and numerous world availability maps are provided in [7.12] for frequencies of 20,30 and 45 GHz.

By way of explanation, consider a typical example taken from [7.12] as shown in Fig 7.5. It is for a geostationary satellite located at 40°W. The operating frequency is 45 GHz and the link margin is set at 12 dB. Based upon the rain-rate map, elevation angle, detailed attenuation calculations, etc., the fraction of the year that the 12 dB margin is exceeded is calculated as a function of geographical location. The estimated availability over an average year is one minus this value. Contours of constant availability are calculated and plotted as shown. As expected, the availability tends to be worst at the lower elevation angles and where the rain-rate probability is highest.

The initial experience of over one year of operation of the Fleetsat EHF Package (FEP) (20/44 GHz) showed that in rain region D, the availability seemed to be actually better than the predicted. The conclusion seems to be that the earlier predictions are pessimistic.

7.3 FREQUENCY BANDS

The available choices are UHF, SHF and EHF. The higher bands that have been allocated for the military but not utilized should also be evaluated for future systems, but are not commented on here. The UHF band is likely to be fully occupied by national systems, in the time frame of interest. It is assumed that any NATO requirements could be met by sharing those systems. The remaining two bands, SHF and EHF, should both have a role. The basic trade-off will be the enhanced AJ protection of EHF versus the value of installed SHF assets. Since the added AJ protection is of most value to small mobile terminals, it seems natural to let EHF serve the bulk of those needs in the next generation system. The large terminals, are not so disadvantaged in EIRP relative to the jammers as are the small mobile terminals. Their natural role would seem to be continuing to provide the high rate trunk line services for the remainder of their natural lives.

Eventually these high rate services too would transition to the EHF band for reasons of increased AJ and spectrum crowding. The question that will face the planners in the time frame of interest is whether the time is right for that transition. In other words, should one continue to put SHF assets into orbit, which will prolong the transition, or should one begin replacing SHF with high rate EHF in orbit. That decision will probably be economic and the results cannot be anticipated at this time. For the foreseeable future it would seem to be SHF for high rate and EHF for small mobile.

Frequency allocations for military SATCOM and some other military systems up to 40 GHz are given in Table 7.1. These allocations have been agreed among the NATO nations through the Allied Radio Frequency Agency (ARFA). In fact, ARFA has also allocated frequency band for military SATCOM at higher frequencies up to 50 GHz: Table 7.2 summarises the EHF military SATCOM allocations and indicates the status of each, i.e. whether shared or exclusive.

At frequencies above 50 GHz a number of frequency bands have been allocated to SATCOM by the ITU. As yet, however, ARFA has not decided which of these will be "earmarked" for military use. The frequency bands which have been allocated by the ITU to SATCOM at higher frequencies in the range 54-300 GHz are indicated in table 7.3. It should be noted that all these higher-frequency bands are shared between SATCOM and other services.

Use of the EHF SATCOM allocations in the range 50-300 GHz may be feasible in the timescale of interest. Reasons for exploiting these frequencies could include the avoidance of congestion, improved A/J due to the large bandwidths available, and reduced probability of intercept due to even higher levels of atmospheric attenuation than that at 44 GHz. While increased attenuation can be beneficial for LPI, it must not be so great as to require impractical EIRPs from mobile terminals. Thus it will be necessary to exploit "windows" in the absorption characteristic such as those at 94 and 240 GHz (see Fig 7.6). It can be seen from Table 7.3 that SATCOM frequency allocations exist close to these absorption minima.

Frequency bands allocated to inter-satellite links tend to correspond to absorption peaks, making such links less vulnerable to jamming from the ground. However their use will have to be considered carefully vis-a-vis optical links because of the threat from spaceborne jammers.

7.4 INTERFERENCE AND NOISE

7.4.1 General Considerations

A major factor that needs to be considered at an early stage in the planning of any future SATCOM system is that mutual interference between the proposed system and other systems, both SATCOM and non-SATCOM, must be kept to an acceptable level. This factor is particularly important if use of the Geostationary Earth Orbit (GEO)is proposed, since the number of satellites (existing and proposed) in some portions of the geostationary arc-in particular that portion which has good visibility from western Europe- is already reaching saturation point in many of the frequency bands of interest. This is clearly illustrated in Fig 7.10 (a) to (d). It is to the noted that a majority of slot allocation refer to satellites not in orbit but being planned.

For the past 20 years internationally-recognised regulations have been in force to ensure that when a new satellite system is brought into use the danger of unacceptable interference to or from other systems will already have been eliminated. This is achieved by restricting signal levels in accordance with certain rules and, where necessary, negotiation with the administrations responsible for potentially interfering systems. The regulations have been drawn up with a view to making efficient use of the available spectrum while having regard to the practicalities of satellite and earth station design. These principles are set out in detail in a report of the International Radio Consultative Committee (CCIR) [7.13].

A very large propotion of the spectrum available SATCOM has been allocated on a shared basis with other services, in particular the terrestrial fixed and mobile services. Thus the avoidance of interference to or from terrestrial links-for example point to point radio relay-is just as important as the avoidance of mutual interference with other SATCOM systems.

The *definition* of what constitutes an acceptable level of interference will depend on the types of modulation used on both the wanted and the interfering links. Performance criteria for a variety of communication links, both satellite and terrestrial,have been laid down by the CCIR.From these, interference criteria for the different link types have been derived These are contained in a number of CCIR recommendations,both for interference between SATCOM system [7.14] and interference to terrestrial links [7.15-7.17]. In general,interference from all sources *(including internal sources) should be considered,but in many* cases it is adequate to consider only "single-entry" interference from a single, predominant source.

7.4.2 Interference Between Satellite Systems

Possible interference paths between satellite systems are illustrated in Fig. 7.7.There are two distinct cases to take account of. That shown in Fig 7.7 (a) where the two systems have uplink and/or dowlink frequencies in common,and that shown in Fig 7.7 (b) where the uplink frequency band of one system coincides with the downlink frequency band of the other. The situation shown in Fig 7.7 (a) is perhaps the most common and has been relevant for NATO in planning the NATO III and NATO IV SATCOM systems. However the situation shown in Fig 7.7 (b) can also occur. For example in the band 8.025 -8.4 GHz which is used for SATCOM uplinks by NATO but which is also available as a downlink band for earth resources and meteorological satellites.

The procedure to be followed for ensuring compatibility with other systems in the cases shown in Fig 7.7 is laid down by the International Telecommunication Union (ITU) in Articles 11 and 13 of their Radio Regulations [7.18]. The process is illustrated in Fig 7.9. It consists of three phases, advance publication coordination and notification. At least two years before the planned in-service date of the new system, advance information is sent to the Internationl Frequency Registation Board (IFRB) and published in their Weekly Circular. Other administrations who believe the new system may cause mutual interference have four months to respond. In consideration of these responses the IFRB then publishes a Coordination Request, naming those administrations with whom *coordination is deemed necessary* and including the additional data required for coordination. *Bipartite meetings* take place between administrations at which possible interference levels are evaluated in detail and agreements reached as to how these can be made acceptable. When these negotiations have been saticfactorily completed the IFRB is notified. All the evidence from the parties involved is examined, and if the IFRB is satisfiedwith the result details of the proposed new system are entered in the IFRB's Master Register.

A simple method of determining whether or not coordination is necessary is laid down in Appendix 29 of the Radio Regulations. This establishes a 4 % increase in the noise spectral density of the affected system as the citerion. In the case shown in Fig 7.7 (a), the increase is the sum of two components: one from the downlink of the interfering system as received through the sidelobes of the receiving satellite ground terminal (SGT), and one from the uplink of the interfering system, received at the satellite but referred to the receiving SGT taking account of the transponder gain and the downlink path loss. In the case shown *in Fig 7.7 (b), only the interference path between satellites need* normally be considered, but again the increase in system noise is referred to the receiving SGT taking accout of the transponder gain and downlink path loss. In calculating the interfering power density levels a reduction factor of up to 4 may be included to take account of differences in polarization between the two systems.

It is not necessary to know the detailed sidelobe structures of the earth station antennas *in establishing the need for coordination.* A "reference antenna pattern" has been agreed for this purpose [7.19]. The pattern has the form $G(\theta) = 32 - 25 \log \theta$ dBi for angles between 1° and 48° from boresight, and $G = -10$ dBi for θ greater then 48°.

The IFRB's advance publication of a proposed system will contain enough information so that, using the reference antenna pattern, another adminstration can determine whether the 4 % threshold is exceeded and hence if coordination will be necessary.

7.4.3 Interference Between Satellite Systems and Terrestrial Systems

Whereas mutual interference between satellite systems is normally calculated on a "single entry" basis, i.e. where each interference source is assessed in isolation, this is not practicable in the case of interference between satellites and terrestrial radio systems, because of the very large number of transmitters and receivers that may simultaneously fall within the satellite's field of view. Also, an individual terrestrial transmitter will normally be within line-of sight of many satellites on the geostationary arc. The approach to minimising interference in these cases has therefore been to restrict the EIRP of both satellites and terrestrial stations in certain critical directions.

Interference paths between a satellite system and a terrestrial system are illustrated in Fig 7.8.

In the case of potential interference to geostationary satellites from terrestrial transmitters, Article 27 of the Radio Refulations applies. This indicates that where practical terrestrial stations in shared bands below 10 GHz that have a maximum EIRP greater than 35dBW should ensure that their peak signal directions maintain a separation of at least 2° from any point on the geostationary arc. In shared bands between 10 and 15 GHz the restriction applies to stations with a maximum EIRP greater than 45 dBW, and the minimum separation is reduced to 1.5°. No restriction applies in shared bands above 15 GHz.

Terrestrial stations in shared bands below 10 GHz that cannot satisfy the above recommendation can transmit towards the geostationary arc provided their EIRP within 0.5° of the arc does not exceed 47 dBW and that their EIRP at an angle α to the arc does not exceed 47 + 8(α - 0.5) dBW, where α is in the range 0.5° to 1.5°.

Moreover, Article 27 limits the maximum EIRP of terrestrial stations in the shared bands to 55 dBW, and also limits the input power to the antenna to 13 dBW below 10 GHz and 13 dBW above.

Article 27 was modified at the 1979 World Administrative Radio Conference (WARC '79) to include frequencies above 15 GHz.

The reverse situation, where satellites potentially interfere with terrestrial receiving stations, is considered in Article 28 of the Radio Regulations. This specifies the maximum power flux density, measured at the surface of the earth, that is allowable from a satellite seen at a given elevation. Different limits apply in different frequency bands. The limits are summarised in Table 7.4. The primary intention of the regulation is to minimise interference to line-of-sight terrestrial links. However, in the 2.5-2.69 GHz band interference to tropospheric scatter links is also a potential problem. Here each case must be assessed individually, with a view to limiting the interfering signal to a *maximum of -16dBW in any 4 kHz at the receiver input.*

Although the frequencies considered in Article 28 do not extend beyond 19.7 GHz, WARC '79 agreed to adopt the values for 17.7 - 19.7 GHz for specific bands between 31 and 40.5 GHz, as an interim measure. Frequecny bands above 40 GHz are considered in a new CCIR Draft Report [7.20], which recommends limits some 10 to 20 dB higher than those which apply below 40 GHz.

Restrictions also apply to the EIRP of satellite earth stations in the direction of the horizon, with a view to minimising interference to terrestrial links. These restrictions are also set out in Article 28 of the Radio Regulations. The maximum permitted EIRP depends not only on the frequency used but also on the elevation of the horizon as measured from the SGT site. Where the elevation is *zero or negative, the maximum EIRP is* 40 dBW in any 4 kHz in shared bands below 15 GHz, and 64 dBW in any 1 MHz in shared bands above 15 GHz. For elevations θ of the horizon between 0 and 5°, the maximum permitted EIRP is 40 + 3θ dBW in any 4 kHz below 15 GHz and 64 + 3θ dBW in any 1 MHz above 15 GHz. No restrictions apply when the elevation of the horizon exceeds 5°.

There is flexibility in the regulations to the extent that EIRPs up to 10 dB higher than those stated may be allowed. However if the SGT is visible across a national boundary the neighbouring nations agreement will be required in this case.

Despite the above precautions, coordination will still be required on a case by case basis when terrestrial stations have less than a certain *minimum* geographical separation from SGTs. WARC '71 adopted a method of drawing a coordination contour around an SGT such that terretsrial stations within the contour would require coordination whereas those outside would not. The method was subsequently refined for frequencies below 40 GHz and is published as [7.21]. Extension of the method to frequencies above 40 GHz is discussed in [7.20]. It should be noted that at these frequencies the coordination contour is likely to be much smaller because of increased atmospheric losses-less than 2 km radius outside the main beam, and between 50 and 150 km within the beam, depending on the elevation angle.

7.4.4 Summary

Mutual interference between SATCOM *systems and interference* between SATCOM and terrestrial systems places fundamental limitations on the radiation characteristics of satellites and SGTs, and on the locations of satellites around the geoststationary arc. However, internationally-agreed mechanisms are in force to ensure that all new systems are properly coordinated with those already in existence or planned. Standards also exist for the maximum tolerable interference levels for a variety of modulation and access techniques. The regulations are currently being extended to shared frequency bands above 40 GHz.

Observance of the regulations should provide NATO SATCOM with adequate protection from mutual interference problems, at least *in peacetime. However, the protracted nature of the* coordination process and the increasing demand for a very limited resource make early declaration of a future SATCOM system essential, particularly if NATO proposes to continue using the geostationary orbit.

7.5 COUNTERING PROPAGATION EFFECTS

(Adaptive Control of Satellite Resources)

An area which requires detailed study concerns control of the satellite communications network. The tendency of moving to higher frequencies to meet expanding demands either for capacity or antijamming capability, accentuates the importance of the network control. This is considered bellow.

The planned use of the frequencies in the EHF band raises the problem of how to cope with random deep fades (as distinct from those that are predictable such as those occurring at solar conjunctions) in the order of 15-25 dB that typical links will experience due to weather effects. In SATCOM systems operating at frequencies below SHF, this problem is solved by having a fixed link margin not exceeding 6 dB which is considered affordable. This same method cannot obviously be used for systems which experience very deep fades. Site diversity is being proposed for Civil Systems which, however, entails, in addition to extra investment in the ground segment and in satellite resources, the requirement for characterisation of the local climate.

Another technique to overcome the effects of deep fades would involve adaptive on-board distribution of satellite resources in response to channing link powers. This would require suitable fading sensors, on-board processors and/or a duplex control link to network control facilities on the ground, and certain adaptive elements exercising the allocation of satellite resources. Within the last area considerable progress has been made in recent years, resulting in useful multibeam antennas (MBA's) and multimode TWTA's allowing apportionment of the satellite's prime power among different groups of downlinks.

Several different strategies of control can be devised:

- Allocation of satellite resources to affected links in proportion to their fading depths, rather than equal *division of available resources among such links.*
- Levying of satellite resources from unaffected links in proportion to their excess margins, rather than applying identical margin reductions to all unaffected links.
- Combination with a demand assignment protocol to ensure equitable protection of most active links.
- Combination with other types of on board processing (eg. signal demodulation-remodulation) to optimise up and down link transmission parameters separately.

It is recommended that satellite processed resource adaption be investigated as a possible remedy against the serious environmental effects in the EHF band. In addition to suitable satellite technology (eg. MBA and wide dynamic range multimode TWTA's), this approach to network control requires *careful assessment of the impact on overall system aspects like* multiple access methods, orbit sharing and network topology.

7.6 REFERENCES

Artificial Intelligence and Neural Networks are techniques which can be used to control the resources of both satellites and ground terminals thus enabling them to function in a desired manner almost autonomously under varying propagational conditions and jamming [7.22 & 7.23].

[7.1] J.J.Spilker, "Digital Communications by Satellite", Prentice-Hall Inc, Englewood Cliffs, NJ,1977,Sec.6-6

[7.2] W.L.Flock, "Propagation Effects on Satellite Systems at Frequencies Below 10 GHz": A Handlook for Satellite Systems Desing, NASA Reference Publication 1108,1983

[7.3] "Propagation Aspects of Frequency Sharing,Interference and System Diversity",AGARD Conference Proceedings No.332,Issy-les-Moulineaux,France,October,1982

[7.4] E.E. Altshuler, "Millimeter-wave Propagation Handbook", NATO document AC/243-D/916,prepared by RADC 30 May,1984.

[7.5] Jules Aarons, "Propagation characteristics of satellite links" in Electromagnetic Propagation Problems in the Tactical Environment,AGARD Lecture Series 120, May 1982

[7.6] Allen L. Jonhson, UHF/SHF satcom propagation and system desing", in "Special Course on Interaction of Propagation and Digital Transmission Techniques", AGARD Report No.744,October,1986.

[7.7] J.A.Frazier and K.G. Castor, "Viterbi versus sequential decoding in a scintillated satellite commucations channel" Conf Record IEEE NAECON 83 PP 564-571 1983.

[7.8] R.K.Crane "A global model for rain attenuation prediction" Proc IEEE EASCON 78,pp 391-395 IEEE Publication CH1352-4,1978

[7.9] R.K.Crane, "Prediction of allenuation prediction, Proc IEEE EASCON 78, pp 1717-1733, September 1980

[7.10] B. Segal, "High-intensity rainfall statistics for Canada, "Communications Research Centre Report No 1329-E Ottawa November 1979

[7.11] L.M. Schwab and A.J. Simmons "a predictor model for EHF communication satellite system availabilities in the presence of rain" paper 25 in AGARD Conf Proc No 332 October 1982

[7.12] L.M. Schwab, "a predictor model for SHF and EHF Milsatcom system availabilities in the presence of rain, "Lincoln Laboratory Technical Note No 1980-15, February 1980.

[7.13] CCIR Report No. 455-3, "Frequency Sharing between Networks of the Fixed Satellite Service" XVth Plenary Assembly, Geneva 1982, Study Group 4

[7.14] CCIR Recommendation 466-3, "Maximum Permissible Levels of Interference in a Telephone Channel of a Geostationary Satellite Network in the Fixed Satellite Service Employing Frequency Modulation with Frequency-Division Multiplex, Caused by Other Networks *of this Service", XVth Plenary Assembly, Geneva 1982.*

[7.15] CCIR Recommendation 356, "Maximum Allowable Values of Interference from Line-of-Sight Radio Relay Systems in a Telephone Channel of a System in the Fixed Satellite Service Employing FM when Frequency Bands Shared", XVth Plenary Assembly, Geneva 1982.

[7.16] CCIR Recommendation 357, "Maximum Allowable Values of Interference in a Telephone Channel of an Analogue Angle-Modulated Radio-Relay System Sharing Same Frequency Bands as Systems in the Fixed Satellite Service", XVth Plenary Assembly, Geneva 1982

[7.17] CCIR Recommendation 558, "Maximum Allowable Interference from Terrestrial Radio-Relay Links to Systems in the Fixed Satellite Service Employing 8-bit PCM Encoded Telephony and Sharing the Same Frequency Bands", XVth Plenary Assembly, Geneva 1982

[7.18] ITU, Final Acts of the World Administrative Radio Conference, Resolution 3, Geneva 1979

[7.19] CCIR Recommendation 465, "Reference Earth Station Patterns for use in Coordination and Interference Assesment in the Frequency Range from 2 to about 10 GHz", XVth Plenary Assembly, Geneva 1982

[7.20] CCIR Report AB/4-9 (draft)

[7.21] CCIR Report 382-3, "Determination of Coordination Area", XVth Plenary Assembly, Geneva 1982

[7.22] "Artificial Intelligence in Communications", IEEE Communications Magazine, March 1988, Vol 26, No 3

[7.23] "Neural Networks in Communications", IEEE Communications Magazine, November 1989, Vol 27 No 11

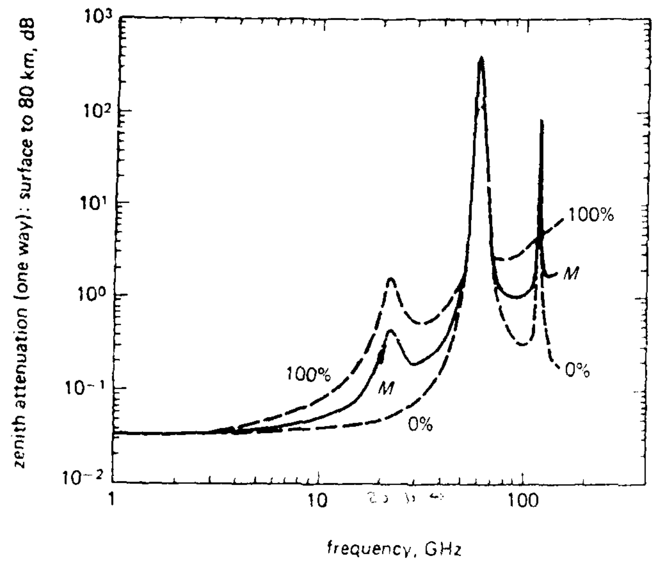


Fig 7.1 Zenith attenuation versus frequency for various percent humidity [7.1]

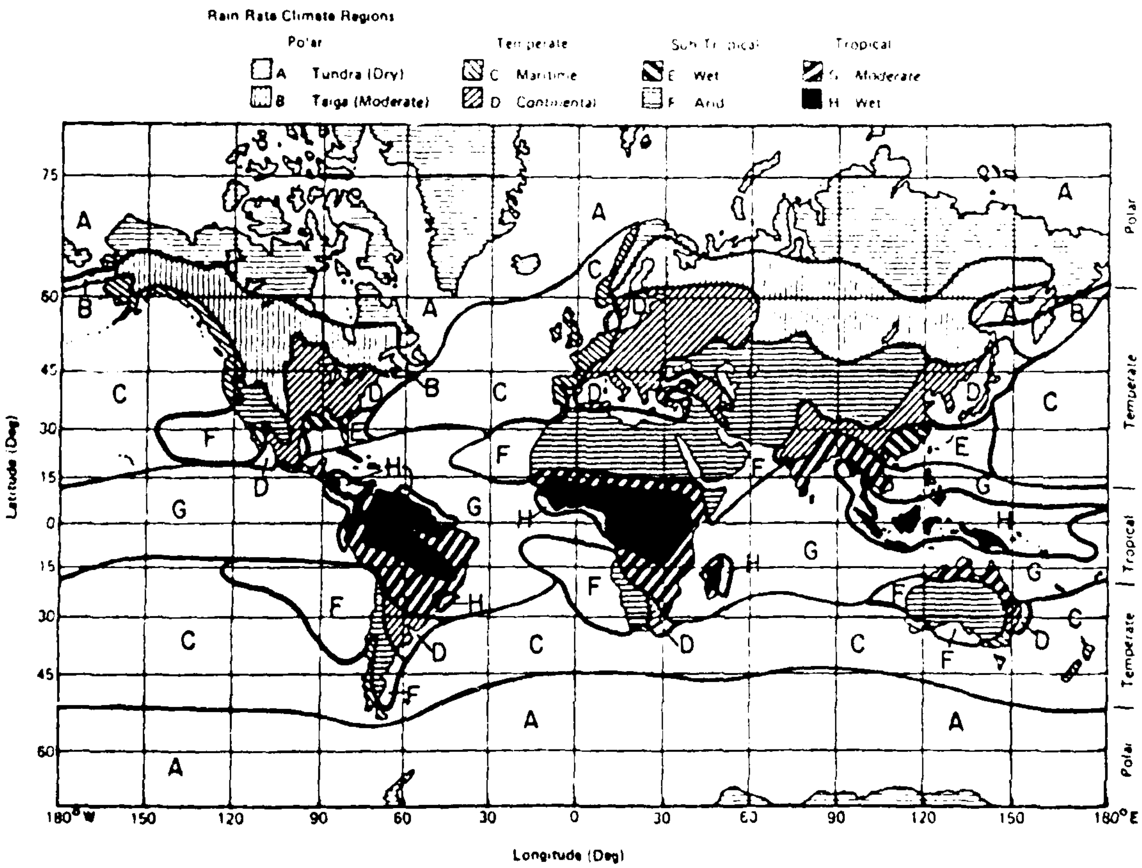


Fig 7.2 Global rain-rate climate regions [7.2]

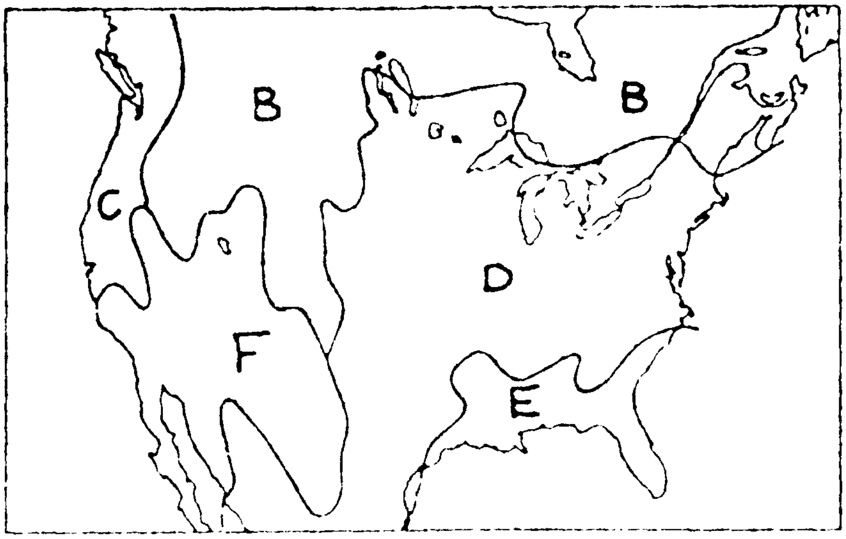


Fig 7.3 Rain-rate regions CONUS and Southern Canada [72]

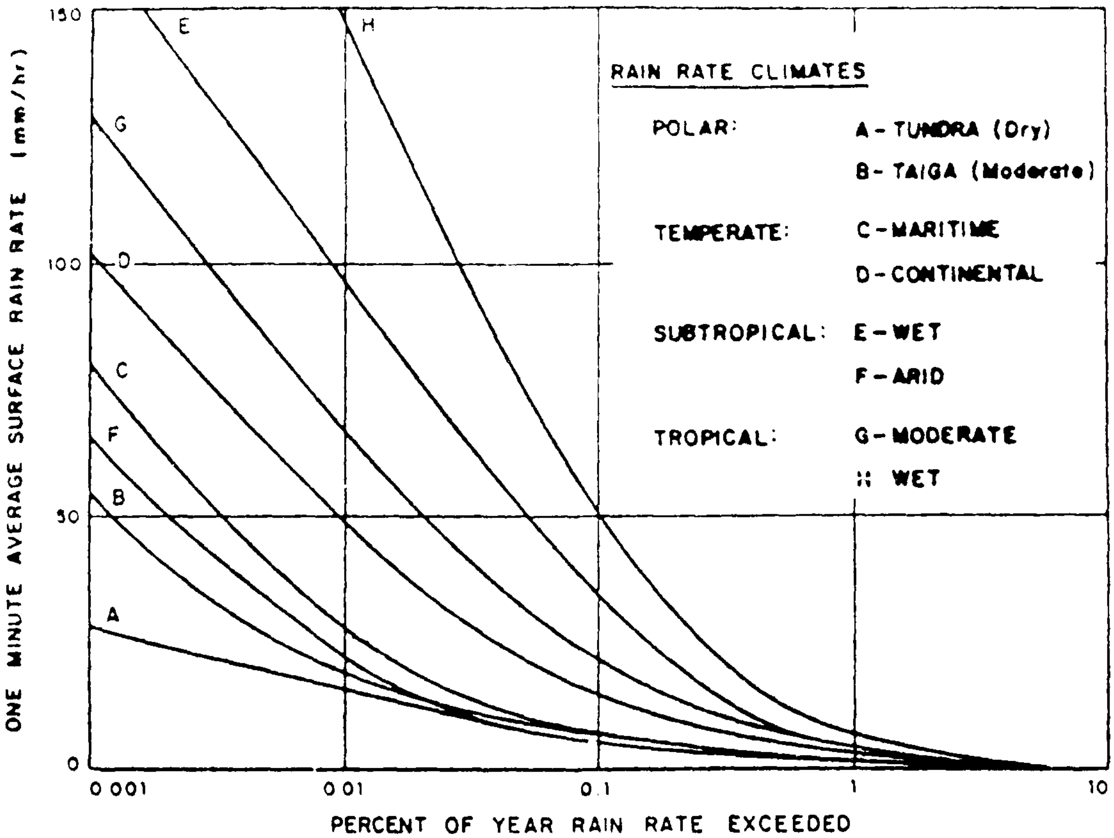


Fig 7.4 Rain-rate distributions for the 8 climate regions [72]

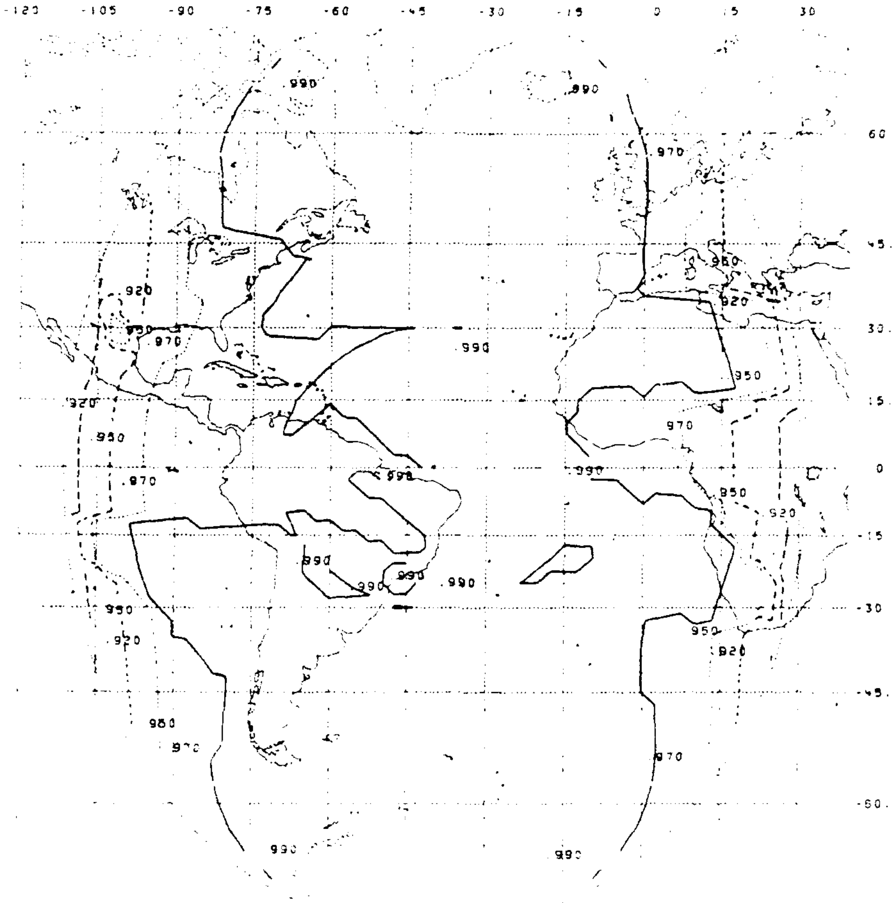


Fig 7.5 FGHZ = 45, DBMRGN = 12, Satellite at −40 deg [7.5]

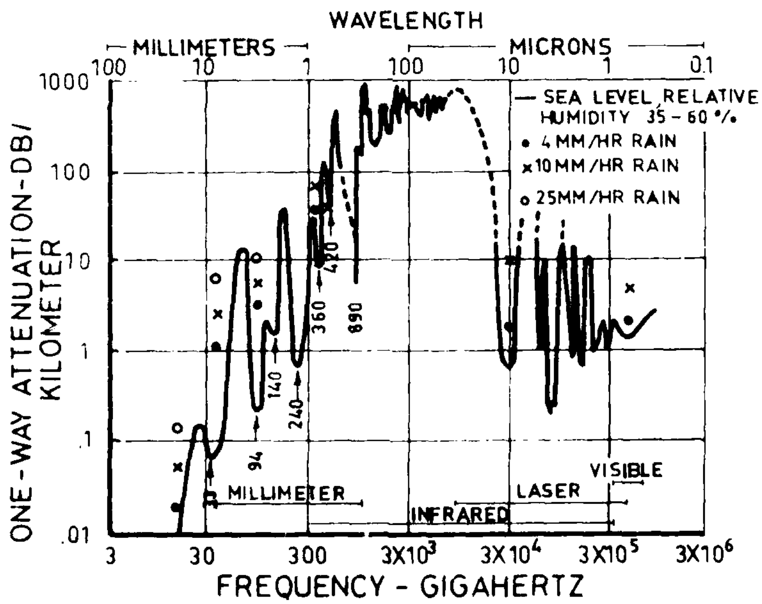


Fig 7.6 Atmospheric attenuation v frequency

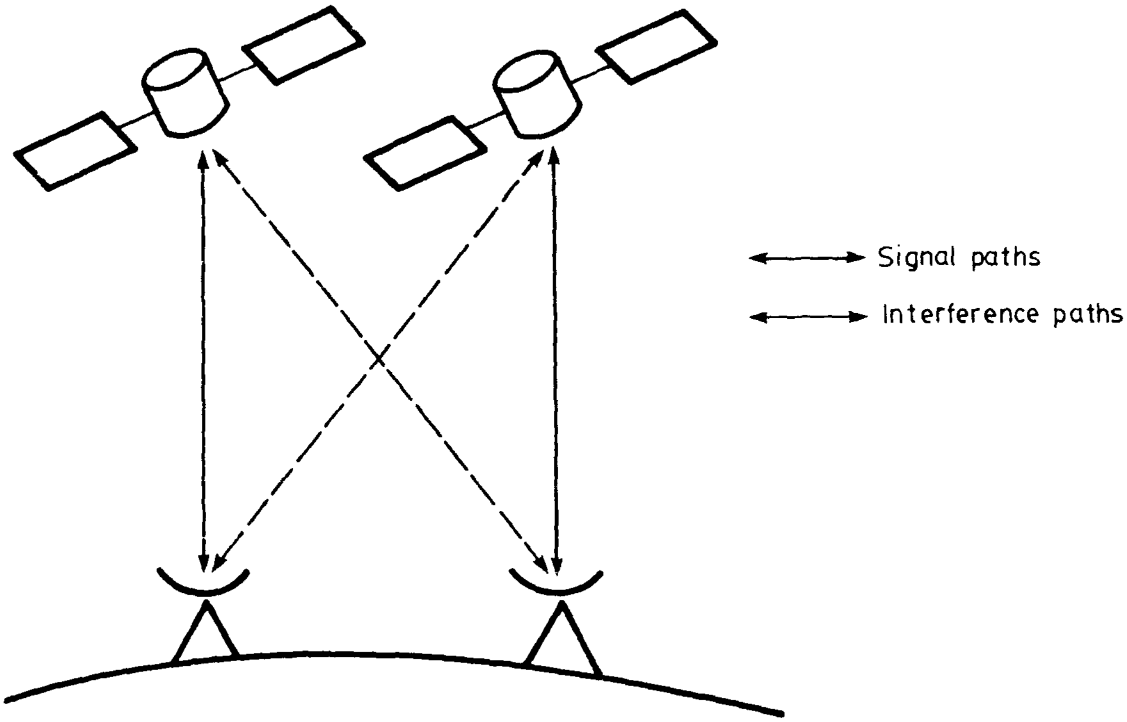


Fig 7.7(a) Satellite systems with common uplink and downlink frequencies

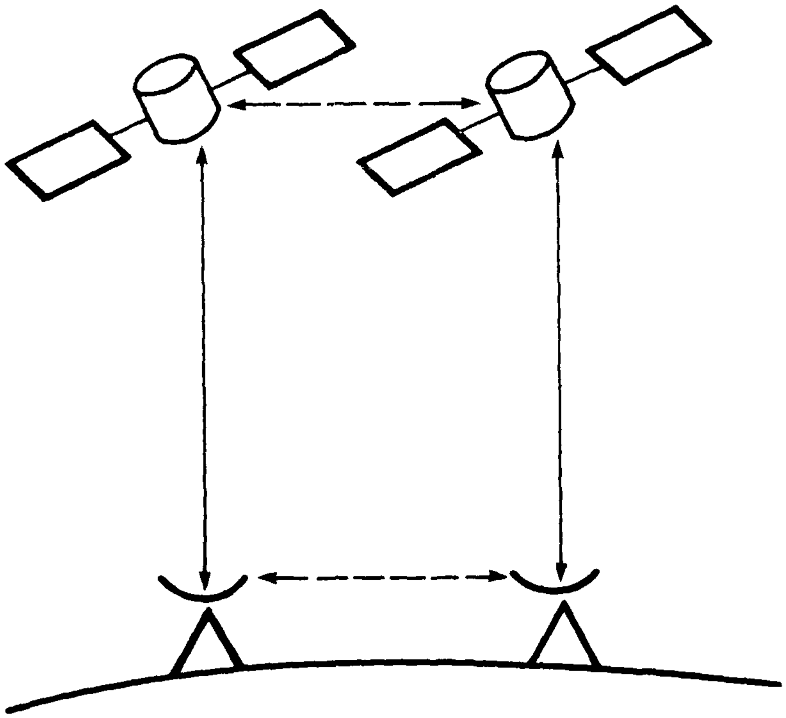


Fig 7.7(b) Satellite systems where the uplink frequency of one coincides with the downlink of the other

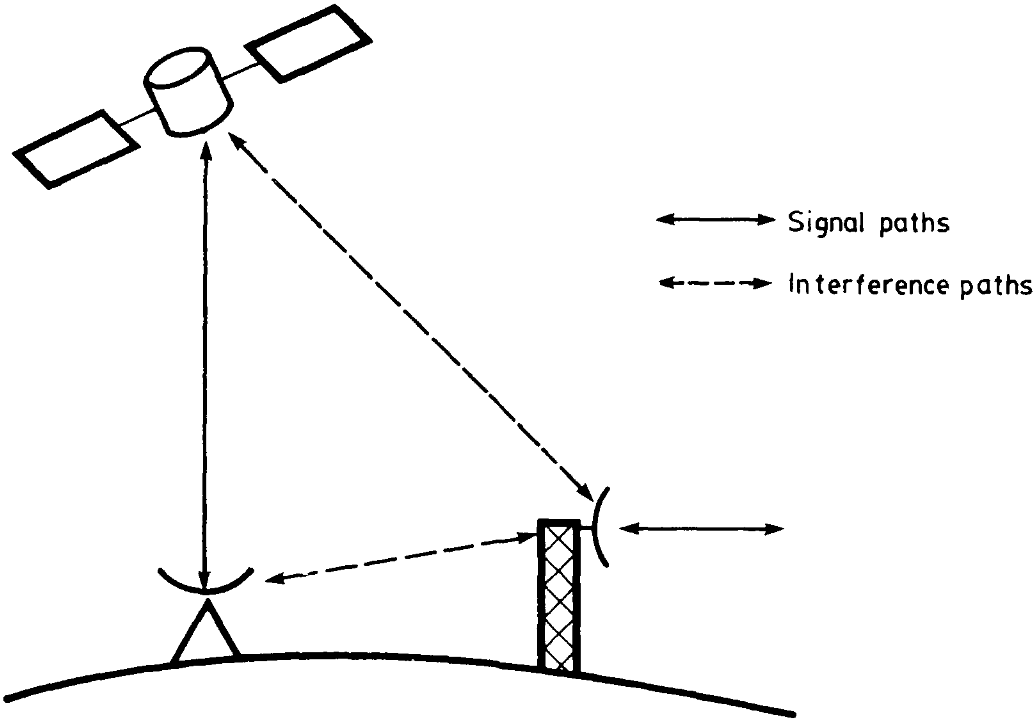


Fig.7.8 Interference between a satellite system and a terrestrial system

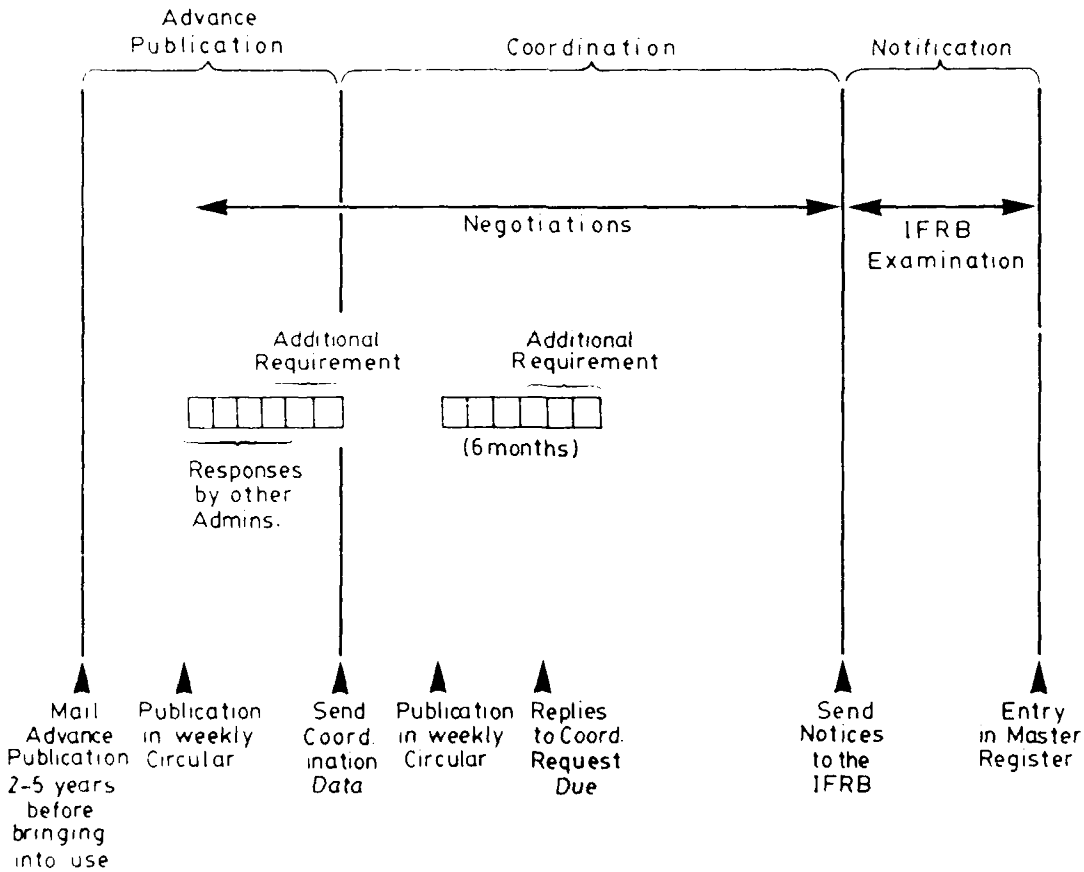


Fig 7.9 Timeframe for ITU process

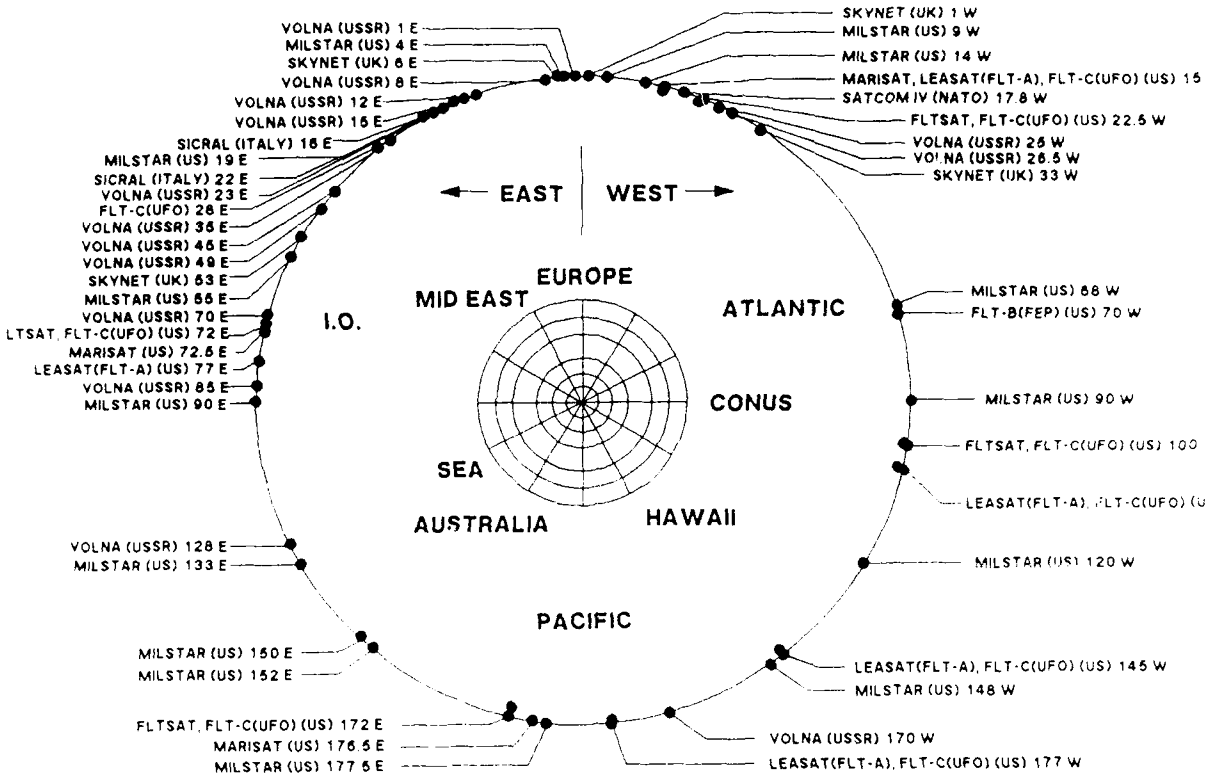
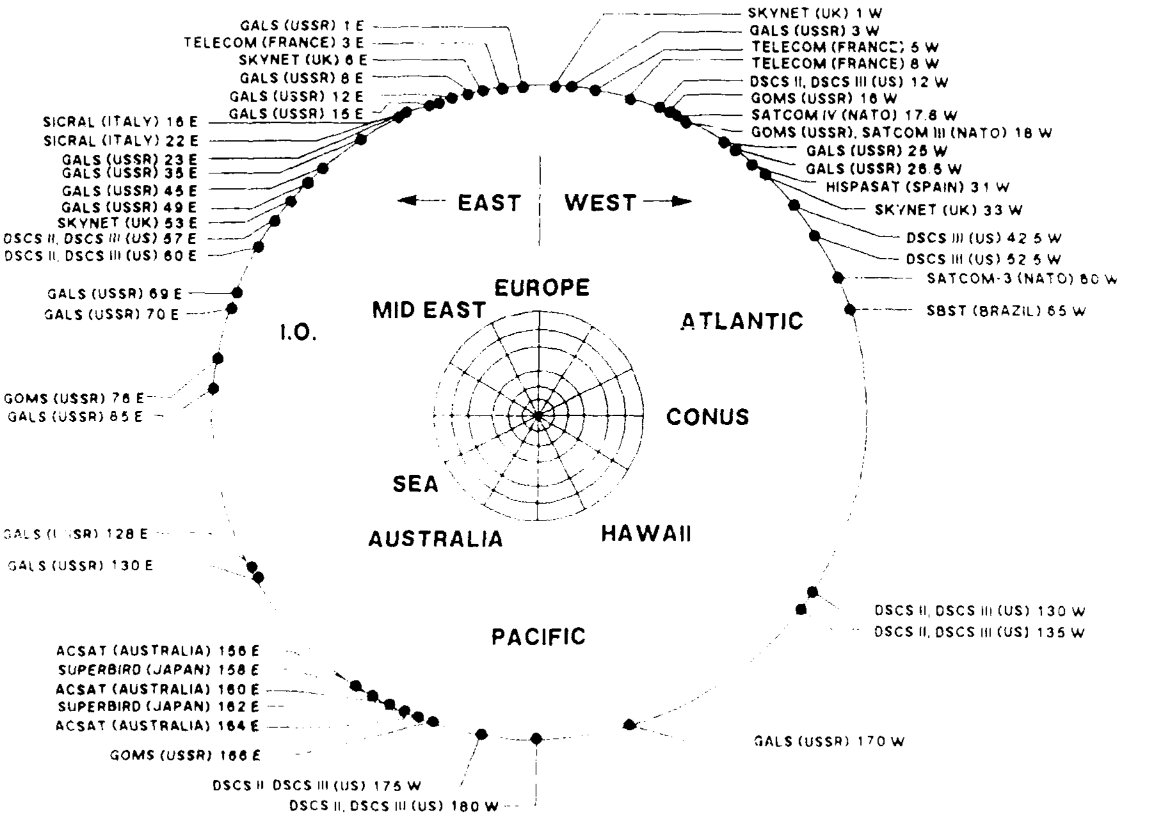


Fig.7.10(a) UHF satellites in geostationary orbit



NOTE FLTSAT SERIES NOT SHOWN (NO TWO WAY COMMUNICATION CAPABILITY)

Fig 7.10(b) SHF satellites in geostationary orbit

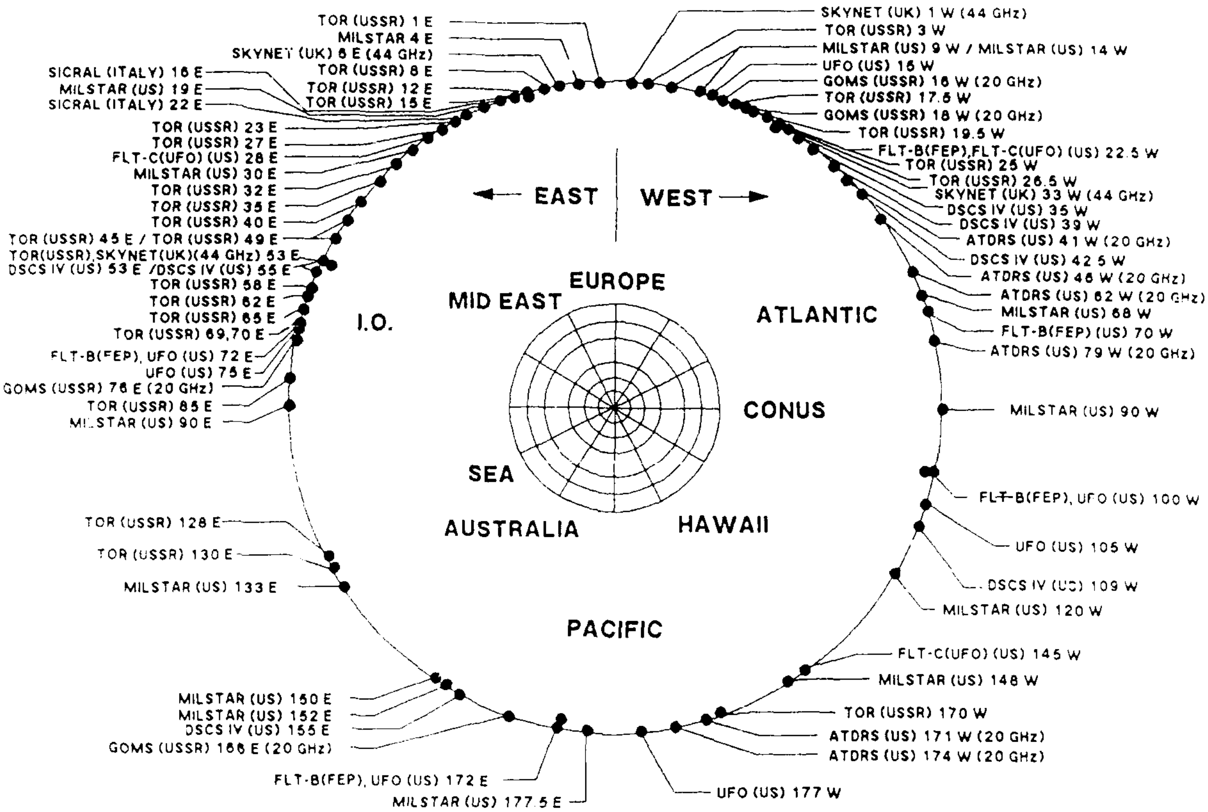


Fig 7.10(c) EHF (44/20 GHz) satellites in geostationary orbit

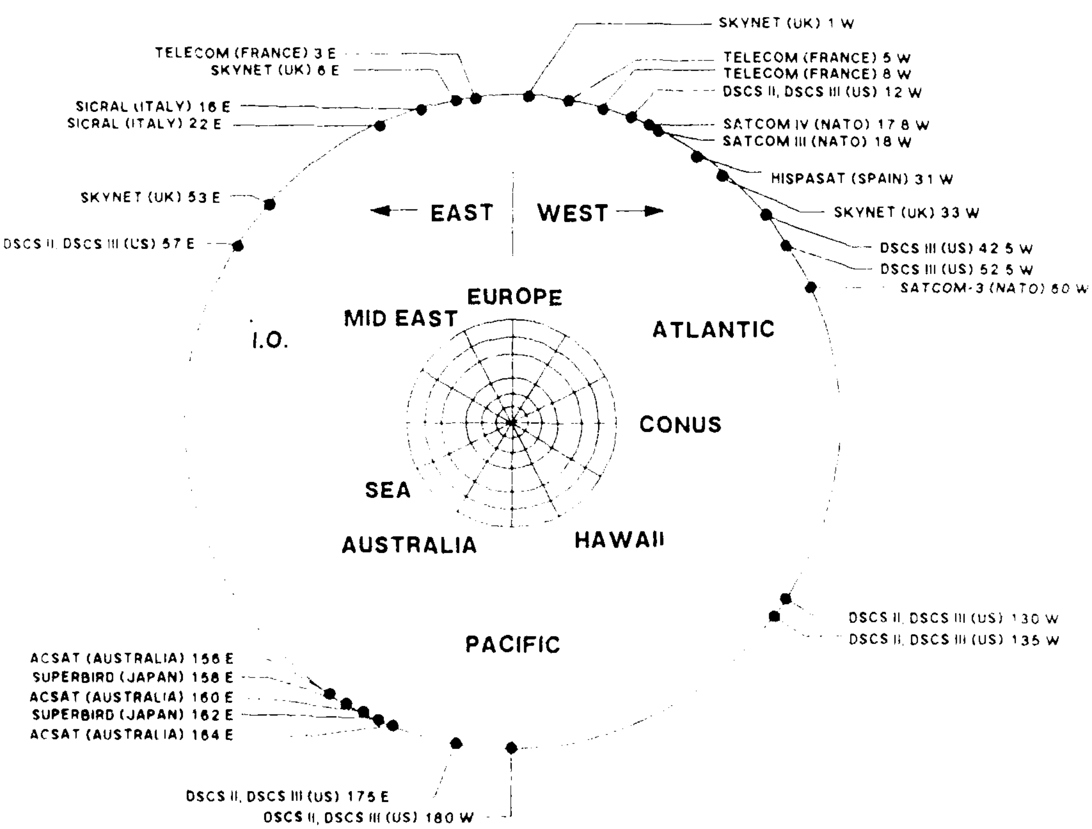


Fig 7.10(d) Allied SHF satellites in geostationary orbit

Table 7.1
Airborne radio and navigation systems and their frequency bands

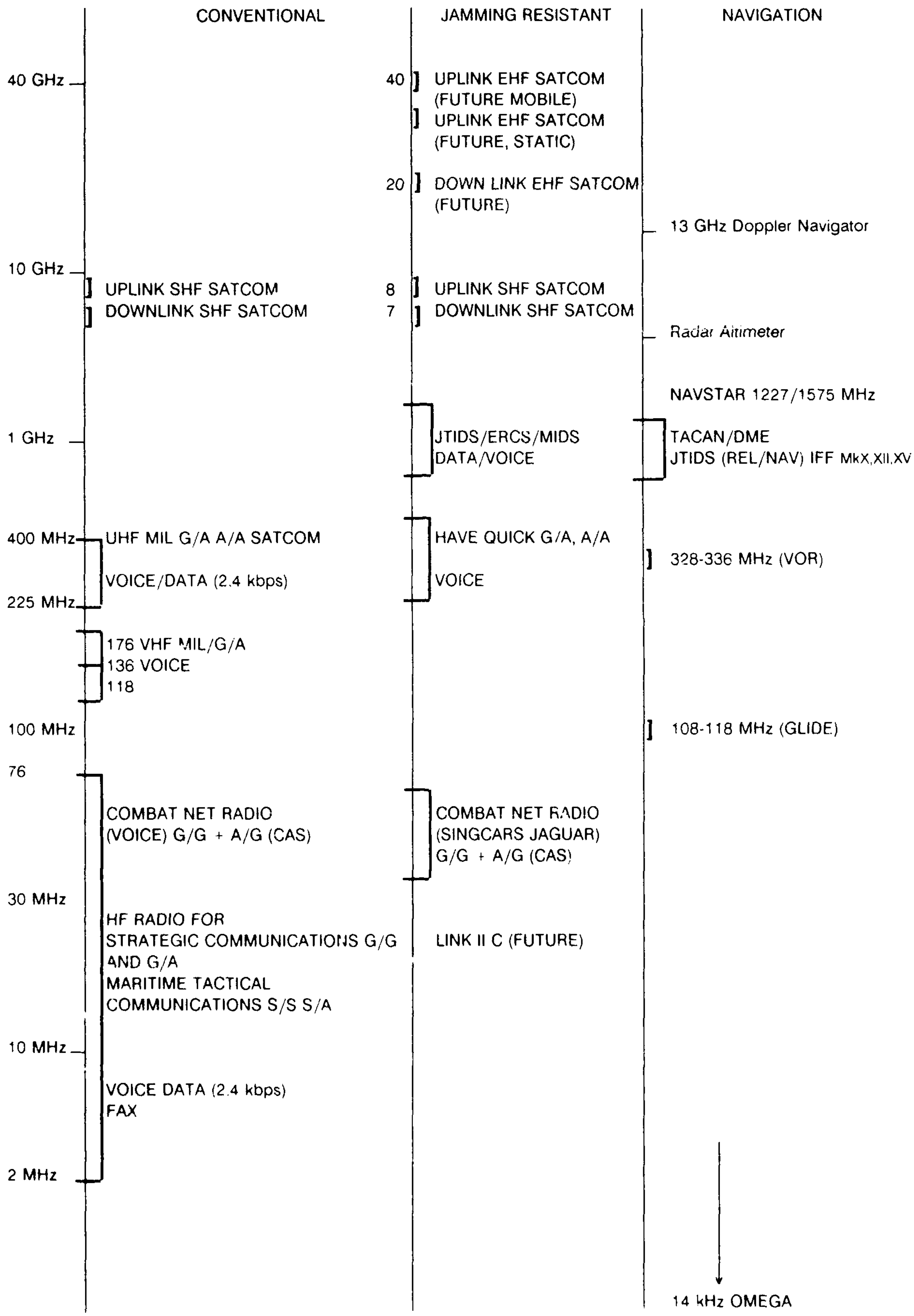


Table 7.2
ARFA allocations for military SATCOM at EHF

Freq. (GHz)	Link Type	Terminal Type	Status
20.2 - 21.0	downlink	Fixed & Mobile	Exclusive military band
30.0 - 31.0	uplink	Fixed & Mobile	Exclusive military band
39.5 - 40.5	downlink	Fixed & Mobile	SATCOM allocation shared with civil
43.5 - 45.5	uplink	Mobile	SATCOM allocation is exclusive military
50.4 - 51.4	uplink	Mobile	SATCOM allocation shared with civil

Table 7.3
SATCOM frequency allocations 54—300 GHz

Freq. (GHz)	Link Type	Terminal Type
54.25 - 58.2	ISL	–
59 - 64	ISL	–
66 - 71	up or down (?)	Fixed & Mobile
74 - 75.5	uplink	Fixed
81 - 84	downlink	Fixed & Mobile
92 - 95	downlink	Fixed
102 - 105	downlink	Fixed
126 - 134	ISL	–
134 - 142	up or down (?)	Mobile
149 - 164	downlink	Fixed
170 - 182	ISL	–
185 - 190	ISL	–
190 - 200	up or down (?)	Mobile
202 - 217	downlink	Fixed
231 - 241	downlink	Fixed
252 - 265	up or down (?)	Mobile
265 - 275	uplink	Fixed

Table 7.4
Maximum satellite power flux density levels in shared bands

Freq. Band (GHz)	Max. Power Flux Density at Earth's Surface
2.50 - 7.75	−152 dB (W/m2) in any 4 kHz
8.025-11.7	−150 dB (W/m2) in any 4 kHz
12.2 - 12.75	−148 dB (W/m2) in any 4 kHz
17.7 - 19.7	−115 dB (W/m2) in any 1 kHz

(i) Satellite elevations below 5°.

Freq. Band (GHz)	Max. Power Flux Density at Earth's Surface
2.50 - 2.69	$-152 + 0.75 (\delta - 5)$ dB (W/m2) in any 4 kHz
3.4 - 7.75	$-152 + 0.50 (\delta - 5)$ dB (W/m2) in any 4 kHz
9.025 - 11.7	$-150 + 0.50 (\delta - 5)$ dB (W/m2) in any 4 kHz
12.2 - 12.75	$-148 + 0.50 (\delta - 5)$ dB (W/m2) in any 4 kHz
17.7 - 19.7	$-115 + 0.50 (\delta - 5)$ dB (W/m2) in any 1 MHz

(ii) Satellite elevation where $5° < \delta < 25°$.

Freq. Band (GHz)	Max. Power Flux Density at Earth's Surface
2.50 - 2.69	−137 dB (W/m2) in any 4 kHz
3.4 - 7.75	−142 dB (W/m2) in any 4 kHz
8.025 - 11.7	−140 dB (W/m2) in any 4 kHz
12.2 - 12.75	−138 dB (W/m2) in any 4 kHz
17.7 - 19.7	−105 dB (W/m2) in any 1 MHz

(iii) Satellite elevations above 25°.

CHAPTER 8

TECHNOLOGY REVIEW

8.1 DEVICE TECHNOLOGIES FOR SIGNAL PROCESSING

8.1.1 Need for New Technologies

Military trends suggest a growing need for complex and comprehensive space borne communication link. To provide such a facility in what is an increasingly hostile operational environment demands a significantly different type of satellite communications system from those currently in service, and changes in the way the space and ground segments of the system are utilised.

The way in which military satellite communications systems will evolve in response to these demands can be viewed at the top level in terms of the type of operational features future systems will need to incorporate to fulfill the role demanded of them. The principal type of developments envisaged are listed below.

a) Higher frequencies of operation with utilisation of the larger available bandwidth for higher rates of data transfer

b) A considerable increase in the use of special waveforms and transmission and coding techniques to minimise the detection interception and disruption of communications by hostile forces.

c) Considerably greater reliability, plus intelligent failure management designed to minimise the operational impact of onboard malfunction and/or damage.

d) A move towards simpler and more mobile ground segments requiring significant increases in the level of satellite autonomy with regard to activities such as orbital maintenance, onboard houskeeping, user location an interfacing, signal regeneration and switching and routing

e) More highly structured systems of satellite with the need for inter-satellite communications for both network management and data routing purposes.

In general these changes will only be made possible by the introduction of new technology. The purpose of the present study is to determine what changes of the sort described above are feasible in terms of available technology and projected technological development, within a time frame extending up to the year 2030.

This particular contribution centers on those technologies which are considered relevant to the fields of analog and digital signal processing.

8.1.2 System Changes

The changes in system utilisation and operational capability that are needed to satisfy changing operational requirements, to improve reliability faults tolerance, to resist active exploitatation and disruption can be expressed as a set of features that a future system would be expected to incorporate. These are summarised below

- Higher freqencies of operation making use of larger available bandwidths for higher data transfer rates.

- New data/voice coding and data compression techniques to make optimal use of available bandwidth

- Increased use of special waveforms and transmission and coding techniques to minimise the detection, interception and disruption of communications by hostile forces.

- Increased use of active antennas to minimise the effects of jamming by employing nulling techniques.

- Data fusion capabilities to reduce pilot workload

- Low level functional redundancy designed into the system which can be utilised during operation to reconfigure the system in the event of failure.

- Intelligent system and failure management capability.

- Modular construction to enable simple maintenence concept to be applied

a) Associated System Functions

In an attempt to translate these operational features into technological terms it is usefull to first consider a possible set of functions for the system being considered. These functions used in conjunction with the predicted trends physical constraints which apply may be used to identify candidate devices and techniques.

Listed below are a set of system functions consistent with the operational features described in the introductory section

a. Initial filtering /preweighting
b. Low noise amplification
c. Antenna switching
d. Hopping/dehopping
e. Frequency conversion
f. Multiplexing /demultiplexing
g. PN spread spectrum processing
h. Digitisation
i. Bit interleaving / deinterleaving
j. Error interleaving coding /decoding
k. Data compression
l. Switching and routing

b) Partitioning of Functions in Terms of Signal and Processing Parameters

Two primary sets of factors which determine the type of technology which is necessary at the various points in the communications signal path are frequency/bandwidth considerations and the degree of speed, complexity and flexibility of the required processing facilities.

The envisaged system would have operational frequencies in the EHF-Band and may wish to utilise the full bandwidth available at these frequencies. In addition advanced antijam and coding techniques, complex routing and highly autonomous operation would demand particularly powerful and in some cases conflicting signal processing needs

High frequency operation presents considerable difficulties at both the device and circuit level because of the problems associated with noise, parasitic effects and the various losses which become significant at these freqencies. Thus early down conversion to an intermediate or baseband frequency is

necessary for efficient performance

The conventional method of down conversion using mixers and frequency synthesisers is still the method generally employed, but work is being carried out in the area of tranverse filters using SAW and CCD device in an effort to develop a means of direct baseband extraction by filtering rather than mixing. The current status of these activities with regard to a possible SATCOM application is not known.

The other priority because of the possibilities it offers for high speed, flexible and controllable signal processing, is to carry out as many of the signal processes as possible in the digital domain. Although it is not certain at which point in the signal path digitisation would in fact occur. It is considered that as listed the processing functions up to (g) would be realised using analog techniques. However category (g) itself represents something of a grey area. It would appear that although *microcircuit technology* is capable of processing spread spectrum signals digitally the upper limit in terms of band spread is relatively low because of the limits which currently apply in terms of speed and resolution to analog – to – digital conversion techniques. Devices in current development offer an upper limit of 12 bit resolution which equates to a band spread of around 200 MHz. This may be inadequate for the sort of bandwidth which will probably be utilised for antijam and LPI purposes by future systems.

Although progress is being made in improving the performance of A to D devices any potential benefits for front end processing has been offset by a continuing movement towards higher frequencies and larger bandwidths, thus the limiting factor in this regard tends to remain. This would indicate that a fully digital communications system is not feasible with technology currently in development and that these systems will certainly require an initial front end and some form of intermediate frequency and analog processing stages for sometime to come.

However recent development indicates that a performance improvement in A to D techniques up to 16-bit resolution is feasible by further development of current techniques and technologies within reasonable circuit bounds. So it is reasonable to assume that the digital boundry will gradually advance and embrace more of the front end functions.

The envisaged system may therefore be regarded as comprising three stages: a high frequency front end, an intermediate frequency and preprocessing stage and a digital signal processing and routing stage.

8.1.3 Analog processing devices

Microminiaturization technologies for analog devices currently considered are e.g.

- Microwave monolithic integrated circuits (MMIC) fabricated in the surface of semi-insulating GaAs wafers utilising linear MESFETs also HEMTs and the heterojunction bipolar transistors (HBT) basing on the use of ALGaAs/GaAs heterostructures.

- Thin film hybrid microwave integrated circuits (MIC) fabricated on the surface of insulating subsrates such as alumina, quarts duraid etc., utilising thin film resistors and capacitors, linear monolithic integrated circuit chips, and passive microwave components interconnected with thin film metal conductor patterns.

- Advanced analog device technologies such as charge coupled devices (CCD) surface acoustic wave devices (SAW) and possibly magneto acoustic devices.

The greatest promise for the type of analog microwave developments listed above being brought to fruition is to fabricate semi-insulating subsrate III-V material using GaAs. Early attempts to construct monolithic microwave integrated circuits or MMICS employed semi-insulating silicon resulting in a very lossy substrates due to the high temperature diffusion processes. The excellent low loss properties of semi-insulating GaAs coupled with its good performance at frequencies above 5 GHz holds significant promise for a true monolithic realisation of CNI system front end.

It is expected, that future systems will combine very-high-speed digital logic and analog microwave front-end circuitry on the same GaAs chip to achieve high-speed signal processing functions.

It may be mentioned that materials such as indium phosphide indium gallium arsenide and to some extent superconductive materials are considered to play a role in the development of high efficiency and low noise components for the use at milimeter wave frequencies.

The area of applications for MMICs includes system functions such as mixers, switches, multiplexers, A/D-military converters power splitters, attenuators, low noise amplifiers, oscillators etc in front-end of many military electronics systems. Significant advantages can be achieved by incorporating passive microwave components such as couplers, SAW filters, combiners dividers, transmit receive switches etc. along with active microwave componenets such as low noise amplifiers, oscillators, IF strips, AGC circuits etc on a common monolithic chip. Improved performance is achieved through the minimisation of interconnect mismatch loses. Interconnection between single-device chips and passive circuitry and adjaining subsrates not only limit performance by reducing bandwidth as a result of associated parasitics, but are labour intensive and less reliable. The monolithic approach integrates interconnections within the chip area and allows only the less critical circuit performance functions to be wired bonded at the periphery of the IC chip. interconnect minimisation reduces cost because wire boding is one of the more time consuming and failure prone operations in the fabrication process. Integration of multifunction circuitry on a common GaAs chip will result in high density microwave microcircuits which will help to balance the effect of ever increasing system complexity in terms of the resulting size, weight, power and costs.

a) Current MMIC Program

Most of the current development in the area of MMIC is being carried out in relation to a US DoD technology program. This was instituted in 1987 when it was discovered that there was little existing development of this type of microwave and military systems operating at these frequencies.

The MMIC program consists of four phases. Three of these phases run consecutively over a planned period of about 7 years (end 1993). Prime contractors of the program are at least 16 US companies. The program aim is to revelop a range of standard monolithic millimeter and microwave devices. So for devices already commercially available, the costs per unit are intended to minimize by establishing more efficient techniques and facilities for large scale production. Other circuits and subsystems which at a more embrionic stage of development will similarly benefit from assistance available from the MMIC program to bring them to maturity. In all cases the emphasis has been placed on the development of those products which are regarded as being most vital to military systems, both for upgrading current systems and those planned for future development.

Gallium arsenide is the primary material employed in the MMIC development program, because it is the most commonly available material that possesses the properties essential for high performance and high efficiency operation at frequencies above 6 GHz.

b) Future Expectation

In the time period 2000-2030 a wider range covering 20-50 GHz can be expected to be available and applications to the transmitter/receiver elements of distributed phased arrays can be assumed. This will lead to reduced size, mass, and power consumption of SATCOM satellites, leading to lower launch costs. The life time of such satellites will also be extended as a result of the greater inherent reliability of the MMIC elements.

8.1.4 Digital Processing Devices

Different digital system functions demand different processing characteristics for their realisation.Function occuping relatively early in the signal path and as such demand high data throughputs for the necessry real time operation.Other functions operating further down stream or not directly on the signal flow may require features that for example enable large quantities of data to be manipulated concurrently but without the same critical speed requirement.

It is normally the case that functions which demand high data throughput rates require structured processing architectures with simple and fixed data paths. In contrast functions which require that complex arithmetic or logic operations be applied to large and possibly varying fields of data need a complex and flexible processing structure in which a measure of speed may be sacrificed to archieve complexity and flexibility.

In the genaral field of processing technology development is being directed to satisfy the particular needs associated with both very high speed processing and cases where processing complexity and flexibility is of more significance. Most of these activities follows from, or are dependent upon current efforts to scale down the feature size of the primary devices. Reducing the overall scale of these primary devices proportionally increases their speed of operation and similary permits more of the devices to be fabricated on a single chip.Thus the benefits of this reduction in feature size is a potential improvement in speed and complexity at a circuit level However this scaling up of the chip introduces latency due to the effective increase in length of some device interconnects so circuit planning becomes more critical.

This reduction in feature size is the principal factor behind the develepment of very large scale integrated circuits (VLSI) and similar activities aimed at producing very high speed integrated (VHSIC) circuits In the latter case the increase in device speed offered by miniaturisation is further enhanced by reducing the potential packing density which is available, so that the devices can be driven harder without breat down occuring and overheating becoming a problem. Higher speed at device level is also offered by the use of new materials the principal one of which is gallium arsenide (GaAs), but these are still comparatively new area of development and the long established materials, in particular silicon, have yet to be matched for all round performance and the clear indicatation is that there is still further improvements in circuit performance to be gained by further development using these materials.

The other areas of activity in the field of digital signal processing are either part of the fundamental process or are only made feasible by virtue of the availability of VLSI and VHSIC microelectronics activities which fall into these categories are listed below

- Microminiaturisation of digital microelectronic integrated circuit technology.
- New design philosophies, methods and computed aided design aids to deal with the vast increases in complexity
- New circuit archtectures.
- New software to take advantage of changing architectures and increasing use of parallel structures to increase processing speed
- New algorithms which simplify and reduce the number of operations required for various common processes

Some of the progress being made in the field of digital microminiaturisation technology is summarised below

a) Microminiaturisation Technologies

- High speed digital monolithic intergrated circuits fabricated in the surface of bulk silicon wafers utilising NMOS,CMOS and bipolar transistor technology
- Very high speed digital monolithic circuits fabricated in the surface of GaAs wafers utilising schottky metal gate field effect transistors MESFETS and Schottky diodes
- Very high speed digital thin film hybrid micro circuits fabricated on insulating substrates utilising the same thin film photolithograph technology as used in the fabrication of MICS. But instead of analog microwave components high speed digital microcircuits which operate in the Gigahertz range are used.

b) Design Trends

There is an increasing use of standard cells techniques in the development of integrated circuit which allows the designer a considerable amount of freedom in optimising circuit placements for a particular application customised chips. Libraries of standart cells are being built up which should give the designer increasing scope to apply this technique. However for military applications where performance is paramount and overheads more of a secondary consideration it is considered that there will continue to be a strong bias towards customised devices.

There is an increasing use being made of CAD/CAE facilities in conjunction with structured design techniques which allows a single engineer to carry out all the tasks required to design a chip and to verify performance using simulation techniques.

c) Processor Architectures

Architecture can be considered in terms of hardware, functional and software structures. The trend would seem to be to align all three aspects more closely so that the fundamental building blocks in all three cases are closely related.

The basic functional architecture uses a single instruction single data stream (SISD) arrangement. This can be enhanced by pipelining which permits the simultaneous execution of certain instructions. Further improvement in processing speed can be gained by using what is refered to as the Harvard architecture which uses dual buses, one for data and one for instructions. This provides a 100 % improvement over tha basic SISD configuration. Reduced instruction set (RISK) machines which provide more on chip registers and eliminate little used instructions as a means of boosting processing speed are a recent development. They can be pipelined to further boost their processing speed. So significant improvements are possible using what is basically a SISD architecture but problems still arise because of the inherent bottleneck which exists.

To overcome the inherent problems associated with SISD topology attention has been turned to various categories of interconnected multiprocessor architectures some of these are listed below.

- Multiprocessor systems with loosely coupled network topologies
- Tightly coupled systolic array topologies
- Multiple process multiple data stream topologies

The introduction of parallel processing structures requires that changes be made to the way in which the system is controlled and also the type of memory provided is centralised, distributed or both.

d) Algorithms

In the general field of digital signal processing the fast fourier transform (FFT) butterfly is still 20 years after conception the computational kernel of current real time signal processing design.

Alternative computational algorithms have been developed and promise improvement in speed and simplicity of implementation by reducing the multiplication burden of standard processing functions (eg convolution) and discrete fourier transforms.

Alternative techniques include the use of:

- The Winograde discrete fourier transform algorithm
- Prime factor algorithms
- Residue number system transforms
- Multiplierless filter algorithms

e) Military VLSI and VHSIC Programs

The DoD 's VLSI and VHSIC programs will include a range of advanced microelectronic circuits for use in future military systems for signal and data processing applications

The programs principle objectives and device parameters goals are listed below.

- Feature size: phase (1) 1.25 micron
  phase (2) 0.5 micron
- Functional throughput:
  phase (1) $5 \cdot 10^{11}$ gate Hz/cm^2 (25 MHz clock)
  phase (2) $1 \cdot 10^{13}$gate Hz/cm^2(100 MHz clock)
- Radiation hard, noise hard, ultra reliability
- minimal size weight, power consumption and life cycle costs
- Test and fault tolarence features
- Minimum non standard VHSIC VLSI parts requirement

At the chip and system architectural level the need for diverse processing types has been identified varying from structured vector and matrix processing to less structured scalar processing. A number of categories of machine organisation have been considered as part of the program for application to structured signal processing and also various pipelining techniques, these include:

* Parallel machines

- Reconfigurable arrays
- Single instruction multiple data flow (SIMD) architectures
- Multiple instruction multiple data flow (MIMD) architectures

* Pipelined machines

- Pipelined vector processors
- Pipelined arithmetic units
- Software pipelining
- Systolic arrays

8.1.5 Development Trends and Limits

Since the the advent of the integrated circuit in 1959, the number of devices that can be squeezed onto a chip has increased from one to several million and as a result the performance of integrated circuits has improved by a factor of around 10 000.

Increasing the raw performance of the basic digital signal processors DSPs is not however the only objective simzlar efforts are being expended to develop complementary technologies such as memory chips, to investigate new architectures and software so as to optimise the processors for the type of real time operation that is now essential and at the same time provide a more flexible device for the user.

In view of the above technological progress the question arises as to how this will continue in the future. To evaluate this it is necessary to determine what factors constrain performance improvement. To do this it is useful to think in terms of a hierachy of limits. Five levels have been identified they are

- fundamental limits,
- material limits,
- device limits,
- circuit limits,
- system limits.

Additionally it is possible to view each level of the hierachy in terms of both theory and practice. Theoretical limits are based on accepted principles of science, whereas practical limits are determined by such things as fabrication processes and equipment.

a) Fundamental Limits

In theoretical terms the fundimental limits are the established laws of physics which by definition are independant of the material employed to construct a device, the design of the device, the circuit in which it is employed and the system in which the circuit is incorporated.

Fundamental limits form the first and most general level of the above hierachy. Fundamental limits come from several areas of physical science. For example, thermodydnamics shows that for any semiconductor device the processed signal has to be significantly stronger than the random statictical fluctuations in the energy which occur at a sub atomic level. These fluctuations impose a minimum switching energy which is temperature dependent. Other fundamental limits stem from quantum mechanics, and from electromagnetic principles or the velocity of light.

b) Material Limits

Material limits depend on the chemical composition and structure of a substance but not on the configuration of the device itself. Several materials are now used to construct semiconductor devices, but silicon is still by far the most common of these. Several properties make it eminently suited for this application, one such property being bandgap which at 1.12 eV is sufficiently large to maintain excellent semiconductor properties over a wide range of temperatures around 300°K. In addition silicon is an abundant elemental semiconductor which can be formed into almost perfect crystals, and its native oxide (SiO_2) is an exellent insulator suitable for integrated circuit fabrication.

The dominance of silicon is now under challange because of a need for higher device switching speed. As a consequence attention is being centered in particular on gallium arsenide (GaAs). In terms of current devices parameters GaAs is about 2.5 times faster than silicon. However as transistors are made smaller this advantage may be offset by other material factors such as thermal conductivity. Silicon having 3 times the thermal conductivity of gallium arsenide can be made into correspondingly smaller devices which would tend to neutralise the inherent speed advantage of GaAs.

c) Device Limits

The predominant device technology in digital electronics is

metal-oxide-silicon. This device is therefore the subject of efforts at miniturisation as a means to boost performance. It can be shown for example that by applying a scaling factor of s, the packing density can be increased by a factor of s*s and the energy consumed in each switching operation can be reduced by a factor of s*s*s.

Given the benefits which follow the scaling down process at device level the question arises as to how far this process can be taken. The limit is directly related to the minimum permissable length of the MOFSET channel, this must be at least as the long as the electron and hole free charge region, which in turn is determined by supply voltage and doping concentration. On the basis of the values which currently apply the minimum theoretical channel length possible is between 0.1 and 0.2 micrometer. Present MOFSET devices have channel lengths between 1 and 2 micrometers.

d) Circuit Limits

At a circuit level the above scaling down process carries possible performance penalties. This is because reduction in device size has been accompanied by an effective increase in the size of chips which in turn increases the length of the chips long distance interconnects. Also although device seperation decreases at a local level there is an attendant increase in current density which can lead to possible interconnect failure due to a phenomenon termed electromigration.

Also of significance at a circuit level is the determination of the minimum supply voltage which supports logic circuit operation. In this respect complementary MOSFET devices require a remarkably low operating voltage of about 0.1 volt at room temperature. A complementary MOFSET configuration also results in the shortest possible channel length of between 0.1 and 0.15 micrometer. These are two of the reasons for the predominance of complementary MOFSET technology.

e) System Limits

The top level of the hierachy of limits is the system level. At this level it is necessary to link all the lower level parameters to the overall system architecture and packaging to determine what the overall limiting factors are at a theoretical level.

From a practical perspective however the limits at this level can be expressed in terms of three parameters. minimum feature size, die area, and packaging efficiency. Minimum feature size is the lateral dimensions of the smallest identifiable feature of a MOFSET or a metal interconnection. A prime example is the gate legth of a MOFSET which is somewhat greater than the length of the channel. There has been a drop in feature size from 25 to 2.5 micrometers over the last 25 years. If this rate of reduction were to continue feature size would be down to 0.25 by the year 2000, which is only slightly greater than the theoretical limits which were said to apply at a device and circuit level. It appears unlikely that this historic rate of decrease will in fact continue if for no other reason than the optical lithographic equipment used to make integrated circuits has begun to reach the limit of resolution which is determined by the shortest wavelength of visible radiation.

Accompanying this reduction in feature size has been a similar increase in die area. This is normally expressed as die edge. Between 1960 and 1980 this has increased from 1.4 mm to 8 mm at this rate by the year 2000 a die edge of 50mm woulld be expected. However practical considerations should modify this rate of increase and a realictic prediction would probably be between 20 and 40 mm.

Packing efficiency as the term suggests describes the number of devices which can be packed on a chip. In the early days of device development progress in this area was rapid. This has subsequently slowed down and recent improvements have only been gained as a result of such things as ingenious design, augmenting the number of steps in the photolithographic process and even building three dimensional devices.

The combined effect of changes in minimum feature size, die area, and packing efficiency has been a rapid growth in the number of devices on a single chip. Despite the limits which have been described and will ultimately determine the maximum number of devices which can be fabricated, it would seeem reasonable to predict that 1 billion devices will be faabricated onto a single chip by the turn of the century.

In view oif the above technological progress the question arises as to how this will continue in the future. Referring to the technology forecast [8.1.16] it is expected that at the turn of the century processor chips (DSP and CPU merged onto one chip) based on BICMOS technology will be available with feature size smaller than 0.5 micron, about 10 million transistors, performing between 100 and 200 MIPS and 300 Mega FLOPs. This will allow the introduction of massive parallel processing systems. Moreover, memory chips are forcasted containing 256 Mbits with access times between 10 and 20 nsec.

8.1.6 Bibliography

[8.1.1] "MIMIC: Improving Military Electronic Systems for the 1990s", E. Cohen, Defence Electronics November 1987

[8.1.2] "C3 and Microwave Technology a Systems Perspective", Walter R. Beam, Microwave Journal December 1986

[8.1.3] "MMIC T/R Modules and Applications", R.Hunter Chiltern, Microwave Journal, September 1987

[8.1.4] "Comparing ADC Architectures is a Designers Best Bet", Mike Koen, Electronic Design January 1987

[8.1.5] "These Chips are Breaking the Linear Bottleneck", Lawrence Curren, Electronics December 1987

[8.1.6] "VHSIC Today and Tomorrow", David O Patterson and E Maynard Jr AGARD Conference Proceedings No 380

[8.1.7] "The Future of Bipolar Technology for Avionics" B Dunbridge, AGARD Conference Proceedings No 380

[8.1.8] "Monolithic Silikon Bipolar Circuits for Gigabit Applications", U Lanmann, AGARD Conference Proceedings No 380

[8.1.9] "ICNIA- Software Programmable Integrated CNI Avionics", Charles W. Ward, 5 th Digital Avionics System Conference Proceedings.

[8.1.10] "Gate Array, Standard Cell, and Fully Custom Building Blocks for the Digital Avionics Designer", H.L. Owen, AIAA Publication 1984.

[8.1.11] "VHSIC Impact on Signal Processing Architecture", T.E. Daly and W. P. Snay, AIAA Publication 1986.

[8.1.12] "Alternative Computer Architectures Reduce Botthenecks", James W Hess, EDN November 1987.

[8.1.13] "Advanced Computer Architectures", G.C Fox and P.C Messina, Scientific American, March 1988.

[8.1.14] "The Impact of Advanced Computer Systems on Avionic Reliability", A.G. Borden, Defence Electronics November 1988.

[8.1.15] "The Impact of Advanced Computational Techniques on Digital Avionic Signal Processing" W.M. Vojir, AIAA Publication 1984.

[8.1.16] "1988 Technology Forecast", Electronic Design, Jan. 1988.

8.2 DIGITAL AND SAW TECHNIQUES FOR ON-BOARD PROCESSING

8.2.1 Processing Principles

Technology progress has made more elaborate on-board signal processing possible. For civilian satellites, on-board processing can be used to enhance the capacity of the satellite by improving link budgets, by adjusting the antenna patterns and by making satellite resources available on a demand basis in (generalized) "switch-board in the sky" concept [8.2.1]. Military satellites has an additional requirement for jamming protection. High-level jammers have two adverse effects on the user signal, (1) the jammer will interfere with the user signal and degrade the total signal-to-noise ratio and (2) the jammer may load the satellites downlink transmitter so that the remainig user's downlink power is insufficient to overcome downlink thermal noise, particular at small SGTs with low antenna gain.

To counter interference (1) above, the users signal is "spread" that is modulated by a code known only to the authorized receivers which spreads the signal over a wide bandwidth compared to the information rate of the signal. The jammer, not knowing the spread code must jam a large number of time-frequency cells thereby giving the user's transmitter a power advantage. The interference effect may be combatted on-ground or on-board the satellite. On-board processing is required, however, to counter effect (2) above. The jammer is rejected in front of the critical non-linear downlink high power amplifier thereby retaining more down-link power for the authorized users. Also on-board processing is based on spread modulation of the uplink signals. Two spread methods are common, (a) direct sequence where the signal is modulated with a pseudo noise sequence known to the receiver and (b) frequency hopping where the carrier frequency is changed usually at a fixed rate but in an irregular pattern known to the receiver. The discussion in the sequel will assume frequency hopping because of its better performance in disturbed and time variant propagation conditions (see Chapter 7) and its better suitability to smal SGTs.

Several combinations of the two methods may be possible. One of the major interest of the Group has been to discuss the evolution of the technological basis for AJ-processing on-board the spacecraft, and how this may make flexible AJ payloads available to NATO in future. NATO has a particular interest in flexible AJ processing because if several national AJ modems could access the on-board processor, NATO SATCOM could develop into a means of AJ communications interoperability between the nations.

The critical components of an AJ system are the AJ receive processor and the means used to obtain and maintain synchronism between the transmitter and the receiver. The spread of the signal over a number of time-frequency cells require synchronism to put the authorized transmitter in a favourable position compared to the jammer. The precision required in synchronism depends on the duration of each cell. For a direct spread system each element of the AJ code (a chip) must be made short to give AJ protection. Frequency hopping allows use of cells with longer duration (the frequency dwell time) and hence less precise synchronization may be used. The implementation of these two functions and methods to enhance flexibility have been considered by the Group as the key topics of on-board processing.

8.2.2 Processor Principles

To receive a frequency hopped signal a bandpass filter with a correct center frequency for each hop is required. The bandwidth of the filter must be suitable to the information bandwidth of the user's signal. The narrower bandwidth, the better AJ (high processing gain) because the number of time-frequency cells is increased due to a finer partioning in frequency.

The sidelobe level of the filter must be compatible with the maximum AJ protection required.

The filter function may be obtained by using one filter with a synthesizer/mixer in front. The spnthesizer is used to translate the received signal into the passband of the filter. It is obvious that this implementation requires one synthesizer/mixer/filter combination for each access. This has no drawbacks for single-users equipment, but it must be expected that a large number of accesses are desireable for a future NATO payload so the amount of on-board equipment may be a limitation to the capacity.

For high capacity it may be more efficient to use a filter bank across the receive bandwidth of the processing payload. In this case a large number of bandpass filters, identical except for their center frequencies are equi-spaced in frequency. To receive a frequency hopped signal the output of the filters must be sampled at points in time in accordance with the hopping pattern. A full filter bank implementation gives a throughput limited by the total receive bandwidth available. Each access has a sequence of filter output sampling unique to that access, however, this function can be implemented mainly in software with a good possibility for reprogramming in orbit and with minor overhead per access. The attractiveness of the filter bank method depends on that a filter bank with N bandpass filters can be implemented much more efficient than N separate filters. This is true both for analog (Sec. 8.2.4) and digital (Sec. 8.2.5) implementation of filter banks. Combinations of de-hoppers and filter banks are of course possible and leads to reduced on-board complexity but also reduction in the maximum number of accesses that can be supported simultaneously. The tradeoff between dehoppers and filter banks will depend on the technology available at the time implementation and NATOs cost/capacity/AJ protection/flexibility tradeoffs.

Fig. 8.2.1 shows a processor according to the filter bank method.

Because the input bandwidth is wide (e.g. 2 GHz) and the final frequency selectivity on the order of kHz, it seems obvious that at least two but most likely three (as shown in 8.2.1) different filter technologies are required, e.g. electromagnetic for coarse filtering, electro mechanic or acoustical for intermediate filtering and digital for fine (high selectivity) filtering.

The resolution of the filter bank may be altered by telecommands during operation to adapt the payload to the jamming level experienced. It is for instance possible to use the high selectivity filter bank only during severe jamming attack and exploit the higher throughput per access possible with the lower selectivity of the intermediate filtering.

The on-board demodulation shown in [8.2.1] has the advantage of completely preventing the power robbing effect (see. Chapter 6). The link budget will also be improved because it gives less degradation to add up-and downlink bit error rates (BER) than to add up-and downlink noise contributions. As an alternative, the uplink filter band outputs may be radiated on the downlink after a suitable frequency conversion which may or may not include downlink frequency hopping. This method gives less restrictions on the modulation of the uplink signals and may therefore be preferable from an interoperability point of view. Also without on-board demodulation a substantial reduction of the power robbing effect may be realized because at each point in time

only a fraction of the receive filter bank outputs will be connected to the downlink transmitter(s).

The Group believes that the flexibility and tha AJ capabilities of future NATO spacecraft will to a large degree be determined by the filter bank technology available and that it is important to NATO to follow, and possibly stimulate, the developments of suitable filter bank technologies during the planning and early procurement phases for future generations of NATO space segments.

8.2.3 Coarse filter banks.

The coarse filter bank operates at microwave frequencies to cover a total bandwidth of 2 GHz. Filter banks at microwave frequencies are implemented today essentially as separate filters connected by non-reciprocal(circulators) or reciprocal (T-junctions) circuits. The performance of the filters is mainly determined by the losses of the cavity resonators used. More than one resonances mode per physical cavity is used to save volume and mass. Higher order resonances are used to reduce losses by increasing the size of the resonator and thus reduce the surface current density. Improvement compared to current technology will be limited unless superconducting materials can be applied, which may lead to dramatic changes in technology level. Superconductivity is addressed in Sec. 8.8.

8.2.4 Intermediate filter bank.

Several technologies may be applied in the future to implement the intermediate filter bank with a selectivity of say 50-200 kHz per individual filter. These include digital, electro-optical, electro-mechanical and surface acoustic waves (SAW).

Based on the current status of these technologies SAW Chirp Fourier Transformers (CFT) processing is considered the most interesting candidate. Digital signal processing (DSP) and the semiconductor technology required is in rapid development and may be a serious contender in the future. At present the modest bandwidth capability of DSP makes it better suited for the fine filter bank.

The Group visited MIT Lincoln Labs in the US and was briefed on the FLEETSATCOM Experimental Package (FEP) and Communication Research Center (CRC) in Canada which also is persuing similar methods for military satellite communication. The Group was also briefed on work sponsored by European Space Agency (ESA) at ELAB-RUNIT and FROBE RADIO A/S both Trondheim, Norway to apply SAW CFT processing to processing payloads for civilian mobile satellite communications. An overview of the results obtained by one of these groups is given in Appendix 8.2.A. The group considers the developments made by these bodies relevent to NATO and the CFT technology a means to implement flexible and robust processing payloads.

Several versions of the CFT is possible, [8.2.2] gives an overview of these. A somewhat simpler version [8.2.3] bandpass filters the signal to be processed, multiplies it with a chirp waveform and convolves the product signal with a chirp in a SAW chirp line as shown in Fig 8.2.2. The mathematical description is given in App. 8.2.A where it is shown that the output signal is given by

$$v_i(t) = \exp(j\Phi(t)) \int s(\tau) w_i(t-\tau) \exp(jt\tau) d\tau \qquad (8.1)$$

where $\Phi(t)$ is a fixed phase function without principal importance
$S(t)$ the receiver uplink signal
$w_i()$ one or several filter function(s) designed into the CFT hardware.

In eq. 8.1 the center frequencies are suppressed. The physical CFT operates at an IF in the VHF/UHF range. h the output $v_i(O)$ is considered it is seen to correspond to $s(\tau)$ filtered by a filter with impulse response $w_i()$. For other values of t, the input signal is multiplied by the term exp $(j\mu t\tau)$ which acts as an oscillator/mixer with oscillator frequency μt. The SAW CFT therefore corresponds to a bank of frequency off-set superhetero-dyne receivers, where all receivers have the same filter response as indicated in Fig 8.2.2

The weight functions $w_i()$ may be a part of the SAW chirp line response. If only one weight function is required for the application this is preferable because the bandpass functions formed by the impedance matching of the SAW transducer will contribute to reasonable weighting functions and reduce the SAW chirp line insertion loss at the center frequency.

If two or more weighting functions are desired suitable filters may be used external to the SAW chirp line. In this case a more or less constant response over the SAW chirp line itself over the CFT integration frame is a natural choice.

The output signals $v_i(t)$ may be processed by analog means, but for AJ porposes it seems preferable to sample and digitize $v_i()$ at intervals δt. The frequency spacing of the CFT slots is $\mu\delta t$ where δt may be selected for a suitable overlap of the CFT slot frequency responses $W(f) = F(w(t))$ ($F()$ - Fourier Transform).

An intermediate filter bank based on the SAW CFT depends on the component technologies:

a) high precision SAW chirp lines
b) high precision IF amplifiers to recover losses in the SAW devices
c) bi-linear mixers to mix the input signal with two or more chirp signals simultaneously
d) high-speed A/D and D/A converters, typically 8-bits.

Components are available today to build SAW CFT with acceptable performance and bandwidths of 10-30 MHz for on-board AJ processing [8.2.3] and for on-board demodulation of civilian mobile type satellite communication. A CFT based demodulator for this purpose with narrower bandwidth has already been demonstrated [8.2.4].

Future NATO space segment will take advantage of a continued progress in analog and digital electronics and it is expected that components (b), (c) and (d) above will be available to build CFTs with bandwidths on the order of 100 MHz or more.

The SAW chirp line will probably remain the critical component with respect to precission ant bandwidth also in the future.

The processed bandwidth will be in the range 30-45 % of the SAW filter bandwidth and the CFT slot selectivity determined by the impulse response duration of the SAW chirp line. These relations lead to a number of orthogonal slot across the processed band equal to approximately 10 % of the time bandwidth product (TBP) of the SAW chirp line.

SAW chirp lines reported so far have the following limitations

(1) T_w < 120 μs [8.2.5]
(2) BW < 1.1 GHz [8.2.6]
(3) TBP < 10000 [8.2.6]

which gives the following limitations for AJ processing

(1) Slot BW min. 33 kHz
(2) Processed band max 450 MHz
(3) Number of slots max 1000

The expected capabilities of future DSP make very narrow CFT slots less interesting, for instance coult 200 kHz slots across 200 MHz processed band be a suitable choice. From extrapolation of experiments. the 10 CFTs required to cover the 2 GHz uplink military SATCOM band would have a total power consumption

below 50 W [8.2.4]. The higher TBPs demonstrated depend on use of lithiumniobate ($LiNbO_3$) as substrate material. The inherent temperature dependence of this material makes some form of temperature compensation necessary. Suitable measures have been demonstrated [8.2.3]. Quartz [8.2.5] (SiO_2) has much less temperature dependence, but the lower piezoelectric coupling means that a lower TBP must be accepted.

SAW devices will benefit from the high resolution lithography developed for semi-conductor manufacture and some modest improvements on SAW chirp line BW and possibly TBP may be expected.

The best SAW chirp line structure for large TBPs is the reflector array compressor (RAC) shown in Fig 8.2.3. The higher frequencies are reflected close to the transducer, while the change in the reflector array spacing causes the lower frequencies to be reflected from the far end of the reflector array. The RAC can be phase corrected by an adjustable phase plate between the arrays [8.2.7]. Amplitude may be corrected by a CERMET film with adjustable conductivity between the arrays [8.2.7] or by 5 adjusting the groove depth [8.2.5]. The response may be corrected to give a slot isolation on the order of 40 dB [8.2.5].

8.2.5 Digital signal processing.

Each CFT slot may be subdivided by means of DSP. DSP algorithms exist to allow filter banks to be implemented much more efficiently that separate filters.

The polyphase algorithms [8.2.8] allow one digital filter to generate all slots. The algorithm is in principle similar to the SAW CFT. The digital filter processing load is proportional to the input sample rate which again is proportional to the number of channels of a given bandwidth. The digital filter is itself therefore has a fixed processing burden per channel independent on the number of channels contained within each CFT slot. (The processing burden depends on the selectivity requirement and the digital resolution).

In the algorithm which is shown in [8.2.4], the digital filter is distributed in a "poly-phase" manner over the inputs to a Discrete Fourier Transform (DFT). The DFT processing burden comes in addition to the poly-phase filter itselves and causes the total DSP processing burden per channel to increase slowly with the number of channels. An oversampling factor of two way be a good choice for DSP filter banks of this type which make half of the DFT outputs useless.

An interesting special case arises when the input consists of only two channels. For oversampling ratio=2 the four-point DFT required can be implemented without multiplications which results in better efficiency for this case. It is furthermore possible to use a socalled half-band finite impulse response filter to obtain the channel selectivity. For half-band filters every second tap weight (except the center tap) is zero which reduces the processing burden by a factor approaching two for high-selectivity filters. The two-channel version of the poly-phase structure is also referred to as a quadrature mirror filter [8.2.9]. To filter out more than two channels from one CFT slot a tree structure as shown in Fig 8.2.5 may be used. Due to the 4-point DFT and the possibility to use half-band filters, the tree structures are more efficient in terms of the required number of multiplications.

The major processing burden stems from fixed point multiplications of rather short fixed point digital words (eg 8 12 bits). To achieve high efficiency both in terms of energy consumed per multiplication and chip area required are the research goals of a major industry. Progress both in Si and GaAs (HEMT) can be expected. The application can carry the cost of the more sofisticated technologies available at any point in time. Currently available 1.25 μm C-MOS technology requires approximately 1.25 nJ per 8x8 bit multiplication.

The selectivity given by the DSP will only be required during jamming attacks. It may be acceptable to equip only part of the total uplink bandwidth with DSP. The DSP filters must then be "hopped" over the band to receive the hopping carriers. This procedure will reduce the on-board DSP burden without consequences to the AJ resistance. However, partial DSP will, at least in principle reduce the number of accesses which may be processed during jamming.

8.2.6 References

[8.2.1] P.M. Bakken and Arne Ronnekleiv: "SAW-Based Chirp Fourier Transform and its Application to Analogue On-Board Processing", International Journal of Satellite Communications, Vol.7, 283-293 (1989).

[8.2.2] M.A.Jack, P.M. Grant and J.H. Collins: "The Theory Design and Application of Surface Acoustic Wave Fourier-Transform Processors". Proceedings of IEEE, 68,pp. 450-468 (1980).

[8.2.3] M.D. Semprucei: "The First Switchboard in the Sky" An Autonomous Satellite-Based Access/Resource Controller". The Lincoln Laboratory Journal, Vol. 1, No. 1 (1988).

[8.2.4] V.Ringset, P.M. Bakken, A. Ronnekleiv, E. Olsen and G. Bjornstrom, "SAW Technology for Multicarrier Demodulation in Advanced Payloads", European Conference on Satellite Communication -1, Munich 1989.

[8.2.5] Arne Ronnekleiv: "Amplitude and Phase Compensation of RAC-type Chirp Lines on Quartz". Proceedings of the 1988 IEEE Ultrasonics Symposium, 1988.

[8.2.6] C.K. Campbell: "Applications of Surface Acoustic and Shallow Bulk Acoustic Wave Devices", Proceedings of the IEEE, 77,pp. 1453-1484, No. Oct. 1989.

[8.2.7] V.S. Dolat, J.H. Sedlacek and D.J. Ehrlich "Laser Direct Write Compensation of Reflective Array Compressors. IEEE 1987, Ultrasonics Symposium Proceedings. pp. 203-208.

[8.2.8] R.E. Chrochiere and L.R. Rabiner "Multirate Digital Signal Processing". Prentice-Hall, Inc Englewood Cliffs, NJ 07632, ISBN 0-13-605162-6. 1983.

[8.2.9] C.R. Galand, H.J. Nussbaumer: "New Quadrature Filter Banks", IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-32, pp. 522-530 No.3, June 1984.

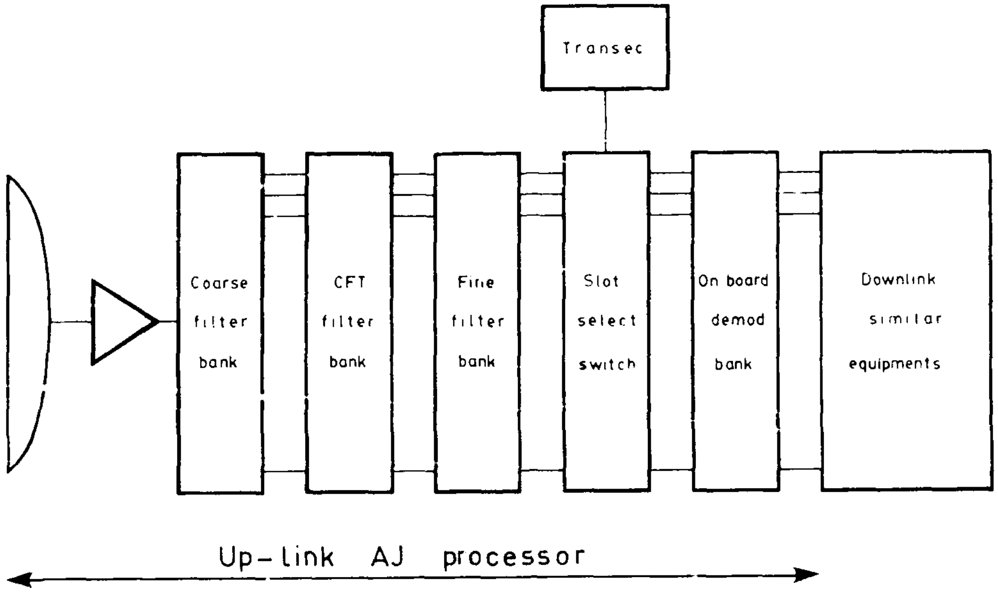


Fig.8.2.1 Block diagram of a processing transponder using filter banks

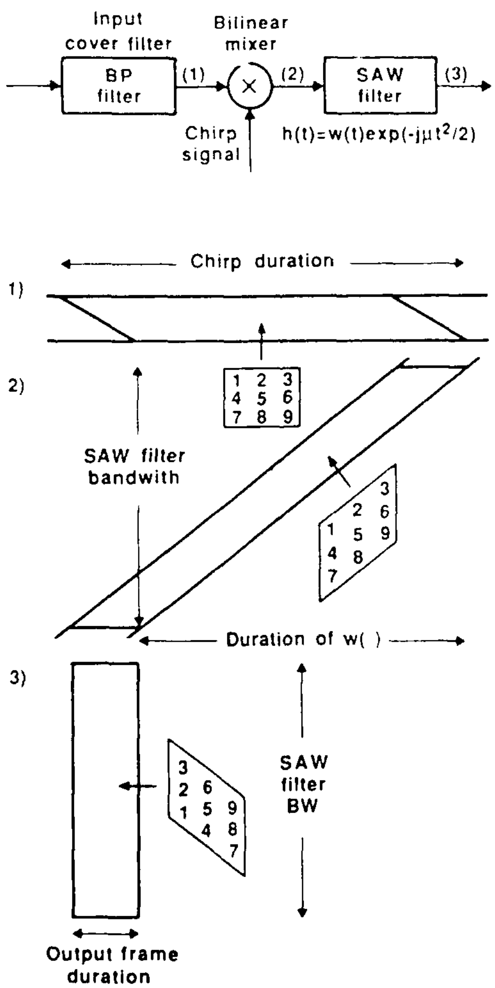


Fig.8.2.2 Chirp Fourier Transformer structure (a) and equivalent signal processing diagram (b). [8.2.1]

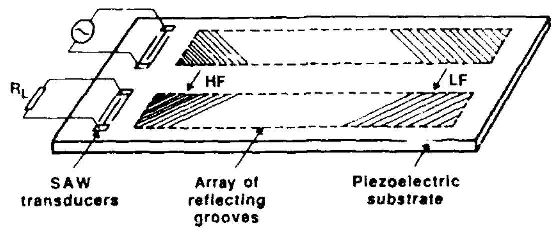


Fig.8.2.3 SAW chirp line with RAC structure [8.2.1]

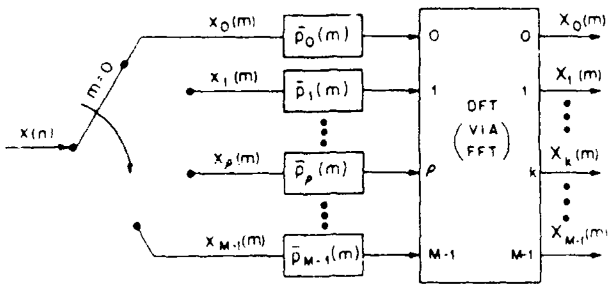


Fig.8.2.4 Poly-phase DSP filter bank [8.2.8]

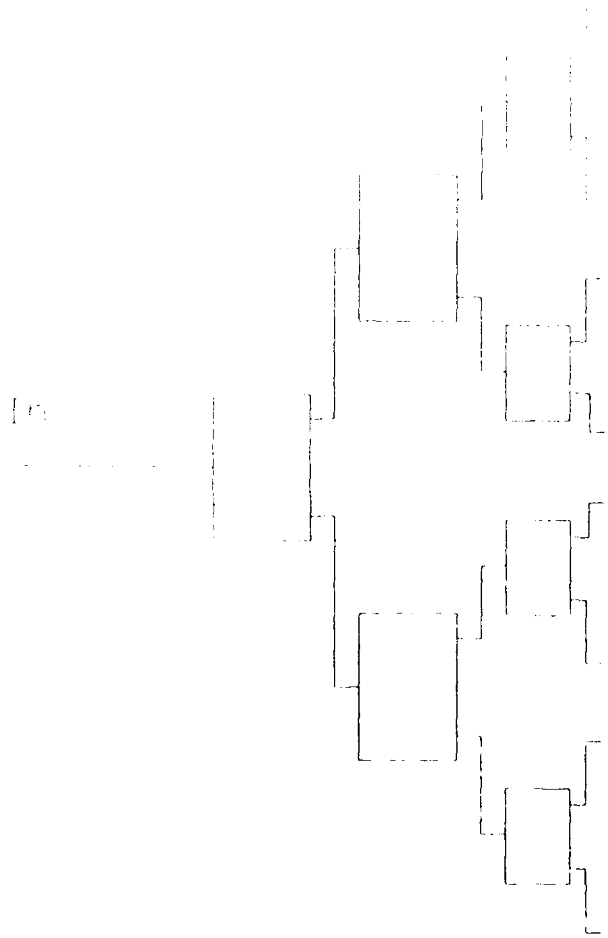


Fig.8.2.5 Tree structure which may use half-band filters to achieve channel selectivity

APPENDIX 8.2 A

SAW - BASED CHIRP FOURIER TRANSFORM AND ITS APPLICATION TO ANALOGUE ON-BOARD SIGNAL PROCESSING(*)

INTRODUCTION

The improvements in optical fibre technology have made it particularly important to develop a technology which enhances the unique properties of satellite communications such as point-to multipoint, communication to mobile users and usage in 'thin' networks where small users are scattered over a large geographical area. The market potential of such applications is strongly dependent on costs, particularly the user equipment cost.

Satellite antennas which generate multiple spot beams provide high reception sensitivity, high effective radiated power and the reuse of the allocated transmit and receive power and the reuse of the allocated transmit and receive frequency band. These are all important contributions to make small, low-cost user stations possible. For an operational system with spot beams the bandwidth and transmit power between the beams must be reallocated during the operational lifetime of the satellite because

(i) the geographical distribution of the users cannot be predicted accurately before launching the satellite.

(ii) it may be required to move the satellite to other orbital locations.

(iii) it may be desirable to introduce new services, or phase out existing services during the satellite's lifetime.

It should be possible to reallocate bandwidth and power with a minimum of bandwidth loss due to the guardbands of the on-board filters.

On-board demodulation offers an attractive, perhaps ideal, solution in terms of guarband losses and reallocation flexibility. However, on-board demodulation may be diffucult to introduce in existing systems, which often have to support a multitude of services and types of carriers including carriers with analogue modulation. To use a number of spot beams to support mobiles requires an on-board flexible frequency selective device which demultiplexes the received up-link signals into one or more sub-bands for each down-link beam. The device must be flexible, i.e. the sub-bands can be altered by telecommands. The importance of narrow guarbands makes it necessary to implement the demultiplexing function at an intermediate frequency and the device is referred to as an IF processor.

It will be demonstrated in this paper that analogue SAW signal processing by Chirp Fourier transforms (CFT) can be used both in on-board multicarrier demodulators (MCD) and in processors for flexible filtering, routeing and beam steering (FROBE). CFTs depend on multiplying and convolving the signal with an adequate precision for digital satellite communications. The most critical components are SAW chirp filters where the impulse response is an amplitude weighted chirp (linear time/frequency relationship). These components have been implemented in our laboratories with an accuracy better than 0.1 dB in amplitude and better than 1 degree in phase (RMS values).

CFT AND ICFT PRINCIPLES

CFT principles

The SAW chirp Fourier transform (CFT) is a Fourier transform valid for a certain bandwidth and over a finite time interval. The realization is based on processing by multiplying and convolving the signal with chirp waveforms.

In the structure shown in Figure 1, the input signal s(t) is multiplied by a chirp

$$c_r(t) = Re[c(t)\exp(j\omega t)] \qquad (1)$$

The centre frequency ω is suppressed in the following description. The chirp signal is therefore described by a complex envelope notation as

$$c(t) = c_e(t)\exp(j\mu t^2/2) \qquad (2)$$

where μ is the chirp rate in rad s^{-2} and $c_e(t)$ is the chirp amplitude. For practical reasons $c_e(t)$ is of finite duration.

The product is then filtered in a chirp filter. The impulse response of the chirp filter is

$$h(t) = w(t)\exp(-j\mu t^2/2) \qquad (3)$$

where w(t) is a weight function which may be tailored to the application. w(t) is of finite duration because the duration is directly related to the physical length of the SAW crystal.

The output of the SAW chirp filter in Figure 1 is

$$u(t) = \exp(-j\mu t^2/2)\int c_e(\tau)w(t-\tau)\exp(j\mu t\tau)\,s(\tau)\,d\tau \qquad (4)$$

The term $\exp(j\mu t\tau)$ in the integral (4) acts as a frequency shift for the input signal $s(\tau)$. At a time t, the output u(t) describes $s(\tau)$ offset in frequency by an amount μt and filtered by a filter with impulse response (w(.) provided $c_e(.)$ is constant over a sufficient interval. The exponential in front of the integral represents a quatratic phase shift. Since the phase shift depends on t, it is in principle known. For the applications to be discussed it is without consequences. The transform may be described as a complete Fourier transform of the CMC type without the first convolution[1] or the sliding chirp Fourier transform.

If u(t) is sampled at $t = t_i$, equation (4) becomes

$$u(t_i) = \exp(j\phi(t_i))\int c_e(\tau)w(t_i-\tau)s(\tau)\exp(j\mu t_i\tau)\,d\tau \qquad (5)$$

which shows that for this sample equation (4) acts as the anti-aliasing filtering structure shown in Figure 2. The equivalent diagram is the familiar frequency conversion-anti-aliasing filter structure. The important difference is that the equivalent oscillator frequency ω_i is determined by the sampling instant t_i. The CFT therefore defines a large number of anti-aliasing filters, centred at different frequencies, using only one physical filter and needing only one A/D converter.

It should be emphasized that the L.O. signal in Figure 2 is not a continuously running L.O. signal, but is started anew for each processed frame, that is for each T_s. This observation implies that the CFT is a sampled system even in absence of an A/D converter.

(*) **Paper published by Petter M. Bakken and Arne Rönnekleiv in International Journal of Satellite Coms. Vol., 7.1989.**

In typical applications, w(.) will be shaped to represent a bandpass filter with a narrow passband compared to the bandwidth of $s(\tau)$. The CFT therefore acts as a filter bank and a means of mapping the input frequency to output time (Fourier transform) at the same time.

The subsequent discussion will assume a CFT transform according to Figure 1.

If two carriers are considered, one at each edge of an input band of width B:

$$s(\tau) = \exp(jB\tau/2) + \exp(-jB\tau/2) \quad (6)$$

two output pulses are produced, centered at times

$$t_1 = B/2\mu \text{ and } t_2 = -B/2\mu \quad (7)$$

and the total time duration is

$$T_a = B/\mu \quad (8)$$

T_a is the active output frame duration and is the time interval on which the input band B is mapped at the output of the CFT. T_a should not be confused with the duration of w( ).

To obtain a finite T_a, the input signal is spectrally limited by a cover bandpass filter which has a finite response in three frequency regions: the lower transition band, the passband and the upper transition band. The two transition bands contain useless signals. At the output these signals are not allowed to interfere with those of the passband, but may overlap in output time with each other as shown in Figure 3. The minimum T_s, after which a new frame may be started is therefore

$$T_s \min = (B_{pass} + B_{trans})/\mu \quad (9)$$

T_s is therefore divided in two parts:

(a) the active frame B_{pass}/μ
(b) the guard time B_{trans}/μ

The active frame relative to the total frame is referred to as the k-factor and is an important design factor. As is evident from Figure 3, the highest k-factor which may be used is

$$k_{max} = 2/(1+s) \quad (10)$$

where is the shape factor, $(B_{pass} + 2B_{trans})/B_{pass}$, of the cover filter. For pratical purposes, s must be defined on the basis of the passband ripple and stopband attenuation which are acceptable for the specific application.

The frame rate $1/T_s$ is also the sampling rate for the output of each filter w( ) of the SAW-CFT-implemented filter bank. If the application requires that all the information contained in $s(\tau)$ is to be represented, the frame rate must meet the requirements of the sampling theorem. To achieve this, the duration of w(.) must be substantially longer than T_s. The consequence is that several chirps must be applied simultaneously to the multiplier of the CFT. The exact number will depend on the application. The use of several parallel chirps is the key to accepting arbitrary symbol timing of the incoming signals. We consider this to be important to the cost of the ground segment.

The duration of w(.) is denoted T_w, and is the integration time for the integral (4). The minimum acceptable chirp signal duration T_c, must allow a full integration time, i.e. overlap between the interval where $c_e(\tau)$ is constant and $w(t-\mu)$ in equation (4), at all points in the active frame or

$$T_c > T_w + kT_c \quad (11)$$

The nominal chirp band is

$$B_c > \mu(T_w + kT_s) \quad (12)$$

The number of parallel chirps is

$$N = T_c/T_s > (T_w + kT_s)T_s \quad (13)$$

The adjacent chirps have envelope functions $C_e(\tau-nT_s)$. These will overlap only partially with $w(t-\tau)$ and cause truncation of w( ). The truncated w( ) has a lower selectivity which may lead to response errors. The countermeasure is to apply a smooth roll-off function for $c_c(.)$ outside the interval T_c.

The signal frequency/time formats at points (1)-(3) are shown in Figure 4. The input to the mixer (1) is bandlimited by the input cover filter. Each chirp will process a limited time window by multiplication into the signal (2). The bandlimiting in the chirp filter reduces the processing time from T_c, the duration of the chirp signal, to T_w, the duration of w( ) (see Figure 4(2)). The convolution in the chirp filter delays high frequencies less than low frequencies, resulting in (3). An interior cell of the processed frequency/time area is also shown. This part of the diagram shows that an FDM-to-TDM transform has occurred and that there is a particular phase shift associated with each channel.

Summary of important CFT properties design choices

$1/T$	The frame rate, which is the sample rate of u(t) for each separate filter in the filter bank, generated by the CFT
T_w	The duration of the integration interval for the CFT. T_w is proportional to the physical length of the SAW crystal and is an important measure of the complexity of the CFT
B_{pass}	The bandwidth of the signal to be processed
B_{trans}	The transition band of the input cover filter

The following major component-related parameters are derived from the design choices

k-factor	The relative part of a frame T_s which contains useful signals $k = B_{pass}/(B_{pass} + B_{trans})$
μ	The chirp rate, positive for the chirp generator and negative for the chirp filter $\mu = B_{pass}/(kT_s)$
B_c	The chirp bandwidth $B_c \geq \mu(T_w + kT_s)$. The chirp bandwidth is an important measure of complexity
B_{SAW}	The bandwidth of the CFT output signal which is equal to the bandwidth of the SAW chirp line. $B_{SAW} = (B_{pass}T_w/kT_s)$. Note that for a well-behaved w(t) the amplitude response of the chirp filter will be close to $w(-\omega/\mu)$ and hence heavily attenuated at the band edges. B_{SAW} is an important measure of the complexity.
BT_{SAW}	The time-bandwidth product of the SAW chirp line. $BT_{SAW} = B_{pass}T^2w/(kT_s)$. BT_{SAW} is a major measure of the complexity for SAW chirp lines.

The inverse CFT (ICFT)

The ICFT may be implemented using the same principles and components as the CFT. In particular, it is important to apply a down-chirp SAW line, because of limitations in SAW technology. If a SAW chirp filter with impulse response $w(t)\exp(-j\mu t^2/2)$ is

considered and an input signal $a_i\ \delta\ (t_i)$ used, the complex conjugated (spectrally inverted) output signal can be written as

$$v^*(t) = a_i^* w(t - t_i) \exp(j\mu t^2/2) \exp(j\mu t_i^2/2)$$

$$\exp(-j\mu t_i t) \qquad (14)$$

The signal $v^*(t)(\,)$ has the desired waveform centred at t_i (if we consider $w(\cdot)$ to be centred round $t = 0$) and three phase terms:

$\exp(j\mu t_i^2/2)$ which is dependent on t_i and may if necessary be corrected for by adjusting the phase of a_i

$\exp(j\mu t^2/2)$ which is dependent on time and must be removed by multiplying $v(\cdot)$ by an up-chirp signal,

$\exp(-j\mu t_i t)$ which is the desired frequency term, translating the output waveform $w(t-t_i)$ to the frequency $\omega_i = -\mu t_i$

By applying equation (14), the ICFT is implemented as shown in Figure 5. Applications of these processing blocks are explained in the 'Applications' section below.

Aliasing in CFT and ICFT structures

As is mentioned above for the CFT case, the mere application of *a periodic multiplying chirp every T_S implies that the CFT is* sampled system. Hence, aliasing may occur between signals which are $2n\pi/T_S$ (n an integer) apart in frequency. Whether aliasing giving rise to harmful outputs does occur in the CFT depends on the length and shape of $w_1(t)$ in relation to T_S, and *on any CFT post-filtering, which may remove aliasing signals.* Post-filtering must operate on signals in a given location in the CFT output time-frame in subsequent frames and, hence, the post-filtering will be periodic in frequency with period $2\pi T_S$. The data filtering in the MCD, and the ICFT filtering in the FROBE structure are examples of post-filtering.

Let us assume that after the CFT signal is subject to a filtering $F(j\omega)$ centred on the CFT equivalent filter centre frequency. If a rejection of aliasing signal components of R is required (compared to the desired signal at $\omega = 0$), the following requirements exists on the CFT-filtering $W_1(j\omega)$:

$$|W_1(j\omega)| \leq R/|F(j\omega)| \text{ for } \omega \geq \pi/T_S \qquad (15)$$

where $W_1(j\omega)$ and $F(j\omega)$ are scaled to 1 at $\omega = 0$.
A possible $F(j\omega)$ and corresponding masking curve for $|W_1(j\omega)|$ is shown in Figure 6. One important fact to notice is that a narrowband $F(j\omega)$ gives more relaxed constraints on $W_1(j\omega)$ from an aliasing point of view. For this reason the required T_W / T_S is smaller for an MCD structure (>2 to 2.5) than for a FROBE system (>4 to 5) for reasonable suppressions.

It should be mentioned that to remove aliasing signals cost efficiently, a frequency-dependent rejection level $R(\omega)$ should be used, because it is easier to reduce the far-off sidelobes of $W_1(j\omega)$.

Aliasing in the ICFT results when signals to be recontructed at ω leaks out on frequencies $\omega \pm 2n\pi/T_S$ (n an integer). Avoiding this restricts the choice of $w_2(t)$ for the ICFT in a similar way to the choice of $w_1(t)$ given above. Any prefiltering of the signals at the ICFT input that may occur may reduce the aliasing problem, and should be taken into account. Since the ICFT input signal is sampled with period $2\pi/T_S$. With a desired aliasing suppression of R, and letting $F(j\omega)$ represent the input spectrum to the ICFT, we find that equation (15) and Figure 6 also apply for $W_2(j\omega)$.

It is possible to calculate the amount of aliasing which occurs in a CFT, ICFT or combined structure. Once the filtering functions $W_1(j\omega)$ and $W_2(j\omega)$ are known.

TECHNOLOGY

The chirp filters for the CFT/ICFT must have a response tailored to the application and better accuracy than that commonly obtained today. This section will give a detailed discussion of the chirp filters.

Other important components are commercially available with specifications in the range which is needed for the SAW FROBE. These are briefly discussed at the end of the section.

SAW chirp line manufacture and correction

Chirps lines can be made in different ways using surface acoustic waves, on a piezoelectric subsrate.[2] For chirp lines with large time-bandwidth products and high accuracy a structure known as the reflecting array compressor (RAC) seems to be the most popular, and is preferred here. A down chirp line is shown in principle in Figure 7, where the input signal is transformed to a surface acoustic wave in the input interdigital transducer and propagates into an array of tilted reflectors in most cases grooves in the subsrate. The distance between the reflectors increases along the array, and when this distance matches the acoustic wavelength of the SAW, reflected waves from several grooves are added in phase, and a strong wave is created propagating towards the other reflecting array. This array in turn reflects the wave towards the output interdigital transducer where it is converted back to an electric signal. In this way signals with different frequencies cross between the arrays at different time delays. A fine adjustment of the time delay or phase with high frequency selectivity may therefore be obtained by spatially changing the SAW velocity in the region between the arrays, *where signals at different frequencies are separeted* geometrically. The velocity change may be obtained in high coupling piezoelectrics ($LiNbO_3$) by shorting the electric fields at the surface by a metal film or by mechanically loading the surface by a film of a material which is different from the crystal. The latter is the only possibility when the *piezoelectric* coupling is low or absent (ST-cut quartz).

For high-precision correction of chirp lines it is important that the correction can be done without changing the device response in uncontrolled ways, such as through wet processing of the crystal surface, which may give contamination, or through changes of strain in the crystal, which may result from demounting and remounting the crystal. This has been achieved on $LiNbO_3$, which has high piezoelectric coupling, by a laser direct-write process[3] for both amplitude and phase correction, or as we describe here for quartz, by sputter etching through a moving slit.

The sputter etching method has ben developed at ELAB/NTH and is described in detail in Reference 4. The amplitude correction part of it only works for lines on quartz. However, since quartz is very temperature stable it is attractive when sufficient bandwidth may be obtained. With this method the ridges between the grooves are covered by a thin chromium film (see Figure 8). The film allows etching of the reflecting grooves and later adjusment of the groove depth to correct the amplitude response. The uncovered quartz is etched by reactive sputter etching, using CF_4 as the etching agent. This does not attack the chromium appreciably. Phase compensation without wet processing is obtained by sputter etching gold from a predeposited gold film between the arrays. The geometric definition of the etched portion of the surface is obtained in both cases by doing the diode r.f. sputter etching through a moving slit in a grounded metal shield, as shown in Figure 9. When the gold film is etched the arrays are covered by thin aluminium strips, and vice versa. When using the slit shown, the actual resolution in the etching, d, is about 5 mm. This resolution is sufficient for almost any correction of the chirp filters with low chirp rates, about 0.06 MHz/μs, which has been made so far. For lines with higher chirp rates μ, etching with better spatial

resolution would be desired, with d ~ 1/√μ. It is believed that the spatial resolution could be improved considerably.

Chirp line accuracy

Both amplitude and phase errors in the chirp lines may degrade the performace of the CFT/ICFT system. The total desired chirp line impulse response is $w_0(t)$ times a phase ...... which is an ideal linear FM chirp. If w(t) is the actual chirp line impulse response envelope obtained, this is related to the ideal response, by

$$w(t) = w_0(t) + \delta w(t) \quad (16)$$

Then a relevant measure of chirp line accuracy is

$$\frac{EE}{SE} = \frac{\int_{-\infty}^{+\infty} \left| \delta W(t) \right|^2 dt}{\int_{-\infty}^{+\infty} \left| W_0(t) \right|^2 dt} = K_{ph}(\sigma\Phi_{w,rms})^2 + K_{dB}(\sigma A_{w,rms})^2 \quad (17)$$

where

$K_{ph} = 3.05 \times 10^{-4}/(^\circ)^2$
$K_{dB} = 1.34 \times 10^{-2}/(dB)^2$

and $\sigma\Phi_{W,rms}$ and $\sigma A_{W,rms}$ are weighted amplitude (dB) and phase errors with the desired response amplitude $w_0(t)$ as weighting.

The ratio EE/SE may be understood as the energy in the signal due to the chirp line errors divided by the energy in the desired signal, when a constant amplitude chirp, matched to the ideal chirp line, is pulse compressed in the line. For satellite communications where carrier power levels are well balanced typical systems would require EE/SE < -32 dB, or chirp line r.m.s. amplitude and phase errors of 0.15 dB and 1 degree, respectively. Spectral components in δw(t) close to an integer times $1/T_s$ are especially harmful.

Processing example at ELAB

Two chirp filters are made and corrected on ST-cut quartz, with a chirp rate ~ 0.0576 MHz/μs, about 40 MHz centre frequency, and an impulse response duration of ~ 120 μs. Only one of the lines with a Kaiser-Bessel type of weighting with ωτ = 7, is reported in detail here. The centre frequency insertion loss of the line is 36 dB. Figure 10 shows the impulse response amplitude, and Figure 11 the deviation from the ideal impulse response amplitude and phase of the filter, both after subtraction of up to quadratic terms in time. For the amplitude, quadratic and linear terms of 0.005 dB and 0.32 dB are subtracted, respectively, over 120 μs. Once these terms are subtracted, the ratio EE/SE is found to be -38.0 dB, or well below the system requirements. The second line was matched to the MCD requirements, and its response is shown in Figure 14. For this line an EE/SE of -39.5 dB was obtained (just after correction; it degraded later due to change of external tuning).

Figure 12 shows the simulated compressed pulse output, with a constant amplitude chirp, matched to the chirp rate of the line as inp it. A direct electromagnetic feedthrough signal, present in the line response is excluded, otherwise no changes are made in the line response. The feedthrough signal may be removed shielding.

Other critical technologies

A highly bilinear four-quadrant mixer is needed to multiply the incoming signal by a number of overlapping chirps in one mixer. Such mixers are commercially available with bandwidths up to 500 MHz which is sufficient for the envisaged on-board processors.

Digital chirp generators require fast D/A converters with low glitch energy. Eight-bit converters in GaAs with speeds up to 600 MHz are now commercially available. The required conversion rate is typically 200-400 MHz.

Eight-bit A/D flash type converters are now available in TTL technologies with speeds of about 100 MSPS for sampling of the CFT output signal. In ECL technoolgy, the speeds are up to 200 MSPS. Our requirement is about three times the processed bandwidth.

APPLICATIONS

Multicarrier demodulation of FDMA carriers

On-board demodulation of FDMA carriers is attractive from a system point of view. The up-link FDMA format allows the ground-station peak e.i.r.p. to be sized to the station bit rate which is important to reduce cost. The demodulated baseband signal in the satellite may be more easily rerouted on- board ('switchboard in the sky') than IF signals, and may also be multiplexed into TDM formats for the down-link which will make it unnecessary to back-off the down-link transmitters.

To implement an efficient on-board multicarrier demodular (MCD), the CFT may be used as a demultiplexer. The CFT is best suited for processing which is independent of the parameters of the incoming carriers. The detailed demodulation can be done by digital signal processing by existing algorithms for symbol timing, phase recovery etc.

Equation (5) shows that for a fixed t_i, the CFT acts as an anti-aliasing filter (AAF) provided that a suitable w( ) is applied.

The centre frequency of the filter is $-\mu t_i$, and by selecting a set t_i across the active frame, a large number (300-1500) of low rate carriers can be demultiplexed. All carriers will share the same CFT, using only one physical receiver, and the same A/D converter. The result is a multicarrier receiver with a complexity similar to a single-carrier receiver.

The sampling rate for each separate carrier at the AAF output port is determined by the chirp repetition pattern no $1/T_s$. We consider it important to apply a sufficiently high sampling rate to demodulate the signals regardless of symbol timing. The solution to this is to use overlapping chirps in the CFT multiplier as explained in the Section on "CFT and ICFT" principles. Because the multiplier chirp is common to a large number of carriers, no synchronism will exist between the sampling pattern and the symbol timing. The digitally implemented demodulator must therefore contain an interpolating filter to produce a data filter output sample at the maximum eye opening.

The maximum selectivity which can be obtained for the CFT depends on the duration of w(.) which again depends on the physical length of the SAW chirp filter of the CFT. The CFT selectivity is an important design parameter which must be traded off against the complexity (number of taps, power consumption, etc). of the digitally implemented data filter in the subsequent demodulator. ELABRUNIT has built [5] a demonstration model (DM) of an MCD for mobile type traffic with 8 kHz channel spacing and 9.6 kb/s data rate. The DM has $(1/T_s)$ = 19.2 kHz or four samples per symbol. The duration of w( ) is 114 μs and the SAW crystal 210 mm. The result is a frequency response with a main lobe of ~32 kHz.

We consider this crystal length to be the maximum practicable and it has been selected to demonstrate to ESA the capabilities of the SAW technology for their comparison of SAW and digital methods for a demultiplexer suited for MCDs. To shorten the crystal, each CFT slot may also be further demultiplexed by

digital processing. For very narrowband carriers (below 9.6 kb/s) additional digital demultiplexing may be the onvious choice. For 9.6 kb/s and higher, combined CFT and digital multiplexing may be preferred after trade-offs between the two technologies.

The MCD DM is built according to the principles shown in Figure 13. The MCD has the properties shown in Table 1.

Table I. Main data for MCD DM CFT

Number of carriers	300
Carrier spacing	8 kHz
w(t) duration	114μs
Chirp rate	57.6 kHz/μs
Chirp repetition frequency($1/T_s$	19.2 kHz
CFT k-factor (active frame/T_s)	0.8

The CFT response W(jω) (Fourier transform of w(t) has been designed taking into account the digitally implemented receiver filter. The response is shown in Figure 14. By design the response W(jω) is minimized in the narrow frequency bands where the digital filters have their repeated passbands. The narrow passband of the data filter makes it possible to use a sample rate which is lower than the width of the main lobe of W(jω). In this way the overall anti-aliasing noise is minimized for the given duration of w(t).

The function w(t) of the chirp line has been realized with an accuracy of 0.08 dB in amplitude and 0.9 degrees in phase (weighted RMS values, see Equation (17)) by using the correction methods described in the preceding section.

The chirp generator has been implemented by a memory of 2048 eight-bit words and D/A conversion. Owing to the periodicity of the chirp pattern only a period corresponding to T_s = 52 μs of the multiple chirps needs to be stored.

For this type of demultiplexer, the major imperfection is adjacent channel leakage caused by the finite duration of w( ) and inaccurate implementation of w( ). The degration due to two adjacent carriers with E_b/N_o = 20 dB has been measured to ~ 0.3 dB. With full loading, i.e. 300 carriers at 17 dB E_b/N_o the degradation due to interference is ~0.4 dB.

The symbol timing has 32 discrete steps per symbol interval which corresponds to a degradation less than 0.03 dB for QPSK filtered with 40 per cent cosine roll-off.

The overlapping chirps may in principle cause degradation due to non-linear intermodulation, however, experience has shown that by careful design no noticeable degradation occurs.

The overall performance of the MCD is identical to high-quality signle single-carrier demodulators. The power consumption will be in the range 10-15 mW per 9.6 kb/s carrier [5,6]

SAW-FROBE

The SAW-FROBE structure is basically a CFT and an ICFT put back to back. With some simple processing between the two, it can provide the following functions:

(a) flexible bandpass filtering
(b) flexible sub-band input-output frequency shift
(c) flexible sub-band routing
(d) flexible beam-forming and steering

The functions all rely on the filtering ability of the CFT+ICFT structure, and hence we first discuss this. The coupling between the CFT and ICFT may be by digitized or analogue samples or by an analogue, gated signal. We first treat the sampled coupling, and then obtain the analogue gated coupling as the limit with infinite sampling rate.

Filtering. Let the output of the CFT be sampled at time instants

$$t_{k,i} = kT_s + \beta_i \tag{18}$$

where T_s is the repetition period of CFT frames and β_i is the subinstant within one frame, $\beta_i < T_s$.

Using that the multiplying chirp signals of the CFT is periodic with period T_s the samples are found to be

$$u(t_{k,i}) = \exp(-j\,\tfrac{\mu}{2}\,(\beta_i^2 + 2kT_s\beta_i)) \int s(\tau)\, w_1\,(t_{k,i} - \tau)\, \exp\,(j\mu\beta_i\tau)\, d\tau \tag{19}$$

Here s(t) is the input signal and $w_1(t)$ the window of the CFT.

Except for a phase term we recognize $u(t_{k,i})$ as the sampled output of a filter $w_1(t)$ centred at the frequency

$$\omega_i = -\mu\beta_i \tag{20}$$

with the input signal s(.) applied. Note that for a filter at the frequency ω_i, an output sample is available every T_s, hence $1/T_s$ is the per-channel sampling rate.

As long as these samples are fed directly into a synchronous ICFT with a window $w_2(t)$ we find the output signal given by

$$s\,(t) = \sum_k \sum_i \int s(\tau)\, w_1(t_{k,i} - \tau) \exp(j\mu\beta_i\tau)\, d\tau\, w_2(t - t_{k,i}) \exp(-j\mu\beta_i t) \tag{21}$$

From this equation we see that for i fixed the $w_2(t)$ window acts as a reconstruction filter at $\omega_i = -\mu\beta_i$ for the elements of the set $\{u(t_{k,i})\}$, without the phase factor outside the integral in equation (19). That the phase factor disappears shows that signals represented by samples taken at different β_i at the CFT output, are reconstructed with the same time delay and constant phase shift at the ICFT output. According to equation (21) the CFT + ICFT structure may be represented by two filter banks, each with one filter centred at ω_i for every β_i, and reconstruction filtering for the per channel sampling rate $1/T_s$, then we see that each switch in Figure 15 controls a frequency-shifted version of W(jω), as shown in Figure 16, which may be used to form passbands and stopbands with intermediate guardbands in a flexible way.

For a simplet interconnection between CFT and ICFT it is natural to let the sampling intervals $\beta_{i+1} - \beta_i = T_A$ correspond to the transition bandwidth, which then becomes μT_A. To obtain a flat passband we must then require W(jω) to be symmetric with antisymmetric flanks, or

$$W(j\omega) = W(-j\omega) \text{ for } |\omega| < \mu T_A$$

$$W(\,(j\mu T_A/2) + j\omega) + W(j\mu T_A/2 - j\omega) = 1 \text{ for } |\omega| < \mu T_A/2 \tag{22}$$

where we assume W(jω) = O for $|\omega| > \mu T_A$.

Figure 17 shows passband ripple and stopband for a combination of four element filters in a simulated case with μT_A = 40 kHz and $T_{w1} = T_{w2}$ = 50 μs. The window functions used are both Kaiser-Bessel windows with $\omega\tau = 7$.

To avoid power consuming sampling between the CFT and ICFT, simpe gating may be used. It should be emphasized that the system is still basically a sampled system, as shown in Figure 15. The main change is that the number of element filters now is

infinite. Hence $w_1(t)$ and $w_2(t)$ must still provide proper anti-aliasing and reconstruction filtering for the sampling rate $1/T_s$. With this provision, and if $p(\beta)$ is a periodic gating function between the CFT and ICFT with period T_s, the total filtering function through the structure is now

$$W_{tot}(j\omega) = W(j\omega) * p\left(-\frac{\omega}{\mu}\right) \tag{23}$$

We hence find that for on/off switching in p(.) the transition bandwidth is now equal to the width of $W(j\omega)$, and that transition bands can be tailored with great freedom by changing p(.). Since there is now no requirement of antisymmetric flanks is now no requirement of antisymmetric flanks in $W(j\omega)$, it may be chosen more freely, and it seems that the connection between window function lengths and minimum transition bandwidth for on/off type of p(.) functions is approximately the same as for the sampled case.

Digital guardband reduction. For the sampled and digitized version of the CFT/ICFT coupling, digital filtering may be introduced to reduce the width of the guardbands. Only $\{u(t_{k,l})\}$ sequences which contributes to the signal in the guarband, that is one or two sequences must be delayed to compensate for the digital filter delays, which will depend on band sharpness and requirements to phase linearity.

Frequency shift. Frequency reuse in a satellite may necessitate frequency translation of narrow or broader frequency bands by different frequency amounts in the IF processor. The SAW-FROBE system allows flexible frequency translation, as the time instant β_l for reconstruction of a signal in the ICFT determines the frequency at which it is reconstructed. Hence, switching the data to the appropriate position in the ICFT input time frame, which normally will require some storage, effects the frequency translation in the digital case. For the analogue case it is more cumbersome, as it requires a signal delay to be introduced in the gating function p(.).

Care should be taken to ensure that the phase term in front of the integral in equation (19) is efficiently cancelled in the reconstruction. It is easily found that this holds for frequency translations $\Delta\omega$ which satisfy the relation

$$\Delta\omega\, T_s = 2n\pi,\ n \text{ an integer} \tag{24}$$

if no phase correction is used. This is a restriction which would exist also in used. This is restriction which would exist also in the equivalent filter bank shown in Figure 15, if it was realized by conventional techniques.

Flexible sub-band routeing. In many applications the satellite operates with several antenna beams of different sizes and orientations. A flexible and efficient use of frequency and power resources can then be obtained by making dynamic reallocation of bandwidth between the various beams.

With one CFT connected to each signal source and on ICFT connected to each output beam, an interconnection between CFTs and ICFTs through a switch matrix is all which is needed to provide full flexibility in sub-band routeing between signal sources and output beams, only subject to guardband and total bandwidth constraints in the CFTs. Figure 18 shows how this may be implemented in the digital case for the mobile communication forward direction. Selected complex signal samples from the source CFTs are added and fed to the ICFTs for each beam.The sample selection is periodic with period $1/T_s$, and hence the routeing information could be stored in rotational shift registers, which are only updated when the routeing is changed. In principle rerouteing of a set oef element filter channels may be done without interrupting the signal flow in any other channels.

One bit is required to define to routeing of each element filter channel. If gain adjustments are desired, this could be obtained by using more bits per element filter command.

Phased array control. In the case where phased arrays are employed as antennas, beam forming may be obtained by feeding the antenna elements, or groups of antenna elements by individually phased and sometimes amplitude scaled signals. For fairly regular beam patterns the required number of antenna element groups is of the same order as the number of beams. Hence, one ICFT may be used for each group of antenna elements, with dynmically changing amplitude and phase control in front of each ICFT. The principle is shown in Figure 19. In this way not only the filtering, but also the antenna beam steering is done in the CFT-ICFT structure. There will be one vector $\{r_i\}$ amplitudes, and $\{\phi_i\}$ phase for each group of antenna elements 1 to J. Each of the vectors will have one element per CFT element filter. Together these vectors control both filtering and antenna beam steering.

This structure is more demanding when it comes to phase stability in the CFT/ICFT structures than the others, as it requires some control of the phase relationship between the outputs from several ICFTs. However, since the phase shifts must be programmable, this may also be used to compensate for phase drifts in electronics or antennas over the satellite lifetime.

CONCLUSION

This paper has discussed on-board signal processing by the analogue chirp Fourier transform (CFT) and its inverse implemented by SAW filters with chirp responses. The CFT/ICFT depends heavily on recent improvements of the precision of SAW chirp lines obtained at ELAB-RUNIT and the progress in semiconductor electronic circuits.

Two applications are discussed in some detail: (1) on-board multicarrier demodulation (MCD) and (2) flexible filtering, routeing and beam steering (FROBE) for transparent payloads. The MCD will have a power consumption of 10-15 mW per 9.6 kb/s QPSK carrier. The FROBE will have a better flexibility to adjust transparent communication transponders after launch to the user requirements than other implementations. Both applications are based on technology which exists today.

REFERENCES

1. M.A. Jack, P.M. Grant and J.H. Collins, "The theory, design and application of surface acoustic wave Fouriertransform processors", Proceeding of the IEEE 68,450-468 (1980).

2. David P. Morgan, "Surface -Wave Devices for Signal Processing", Elsevier, Amsterdam, 1985 Chapter 9.

3. V.S. Dolat, J.H.C. Sedlacek and D.J. Ehrlich, "Laser direct write compensation of reflective array compressors", IEEE 1987 Ultrasonics Symposium Proceedings pp. 203-208.

4. A. Rønnekleiv, "Amplitude and phase compensation of RAC - type chirp lines on quartz", IEEE 1988 Ultrasonics Symposium Proceedings, pp. 169-173.

5. P.M. Bakken, V.Ringset, A. Rønnekleiv and E. Olsen, "Multicarrier demodulator using analog and digital signal processing", Tirrenia International Workshop on Digital Communications, September 1987. Proceedings published by Elsevier, Amsterdam, The Netherlands, 1988.

6. V. Ringset, P.M. Bakken, A. Ronnekleiv, G. Bjørnstrom and E. Olsen, "SAW technology for multicarrier demodulation in advanced payloads", First European Conference on Satellite Communications, Munich, September 1989.

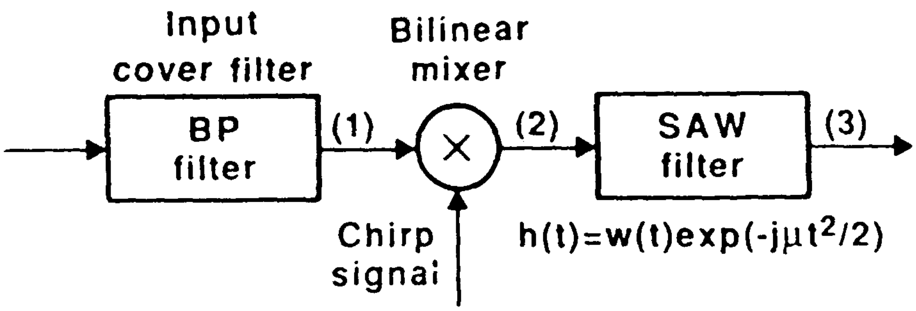


Fig.1 Principle of a CFT according to the MC implementation

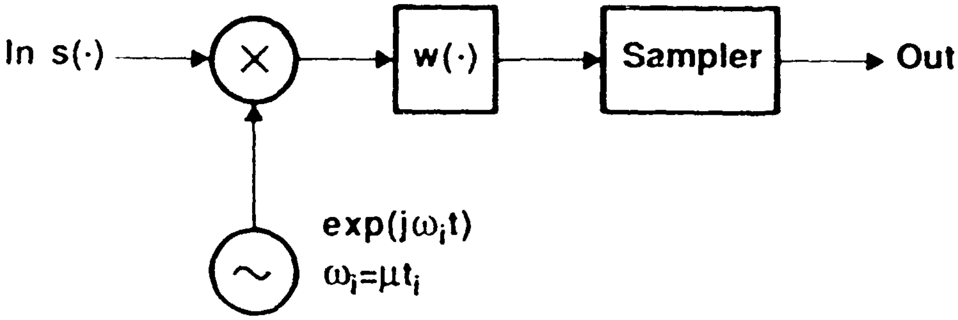


Fig.2 Equivalent diagram for a CFT according to equation (4) with the output signal sampled at $t = t_i$

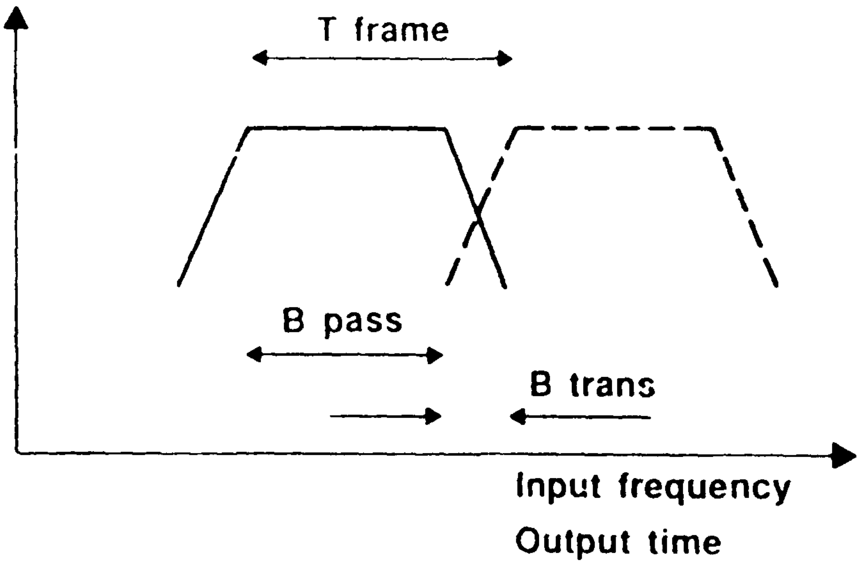


Fig.3 The mapping of the input spectrum onto the output frames

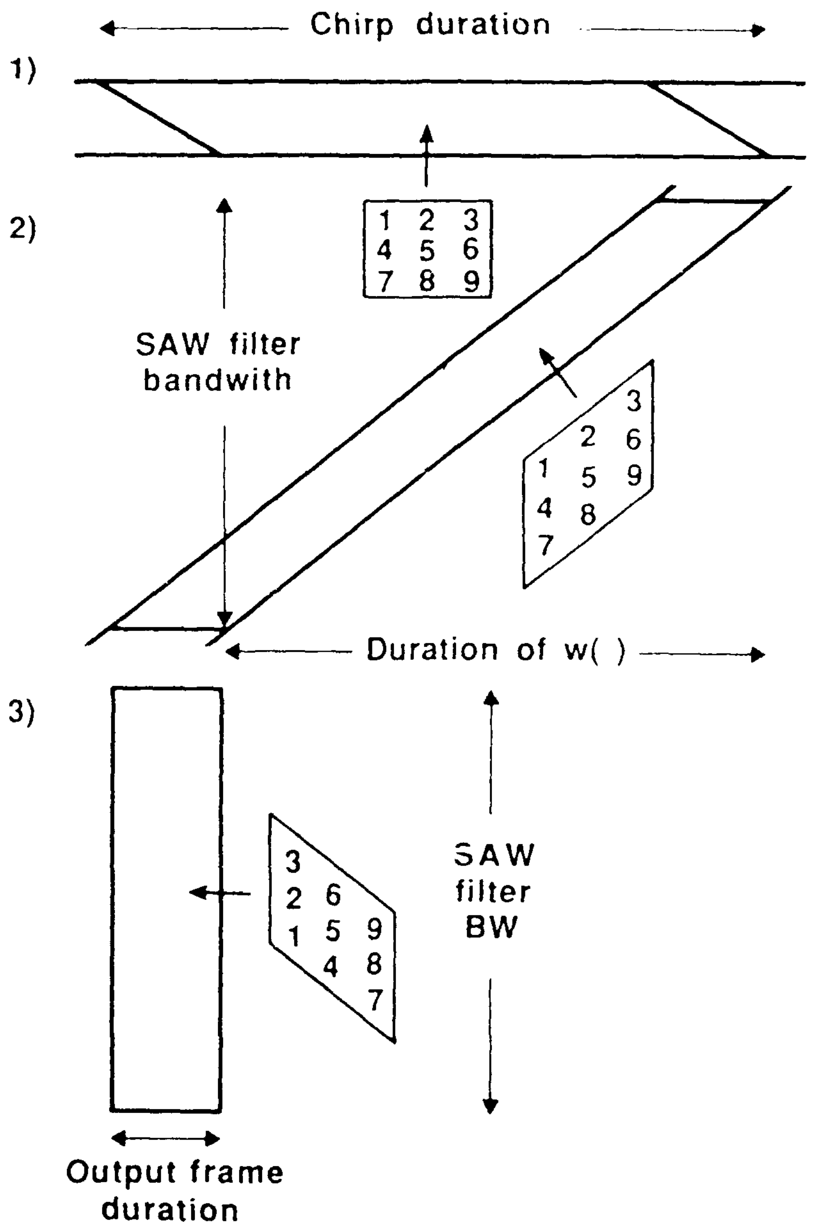


Fig.4 Frequency/time diagrams for the CFT

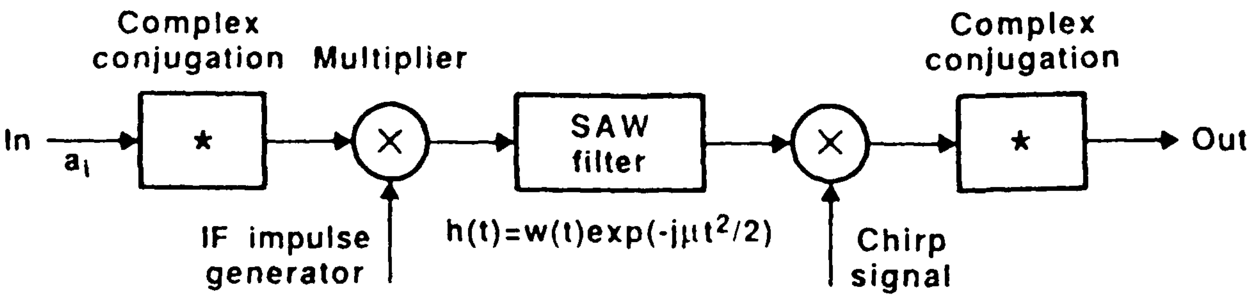


Fig.5 Principle diagram of inverse chirp Fourier transform

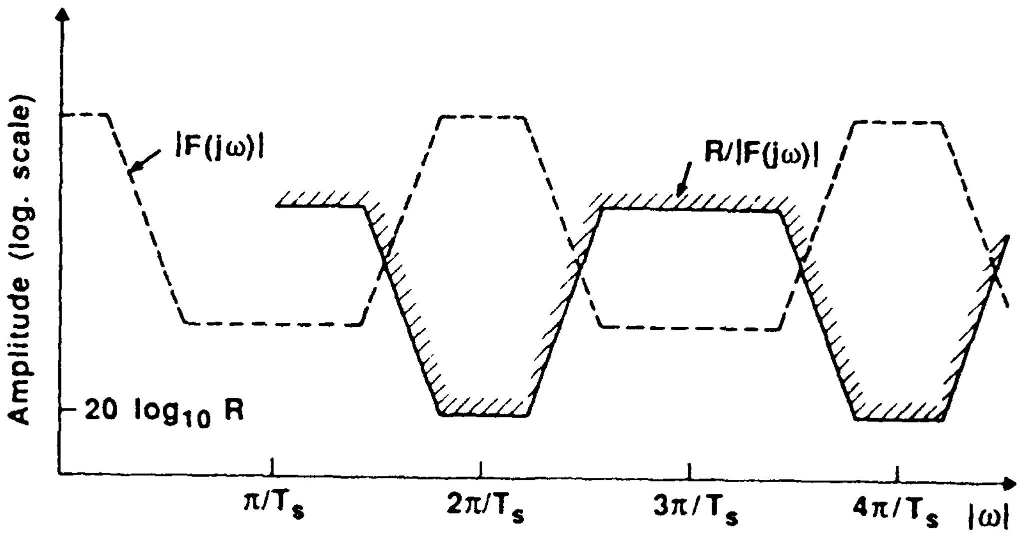


Fig.6 Masking curve (full line) for $W_1(j\omega)$ or $W_2(j\omega)$ to give an aliasing suppression of R. $F(j\omega)$ is post-filtering for the CFT, and the input spectrum for the ICFT

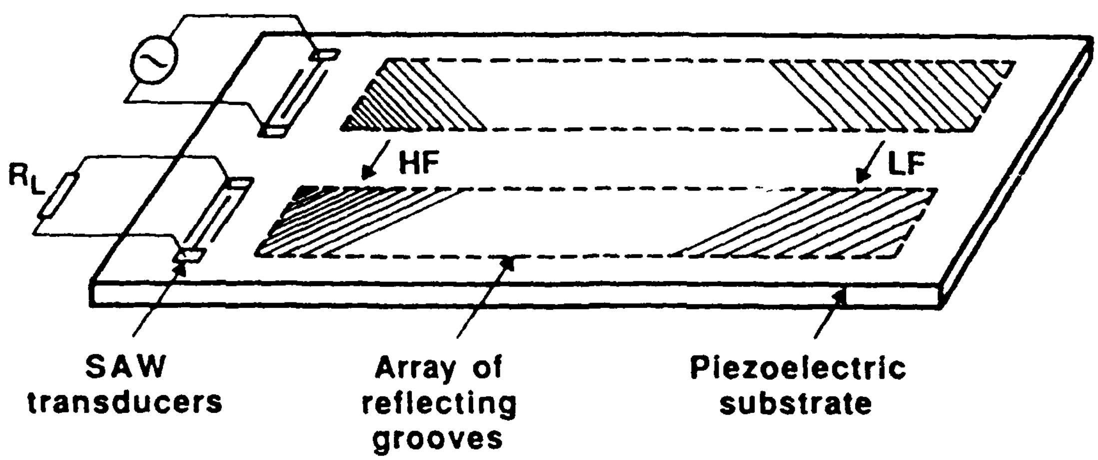


Fig.7 Principal structure of the SAW reflecting array compressor (RAC)

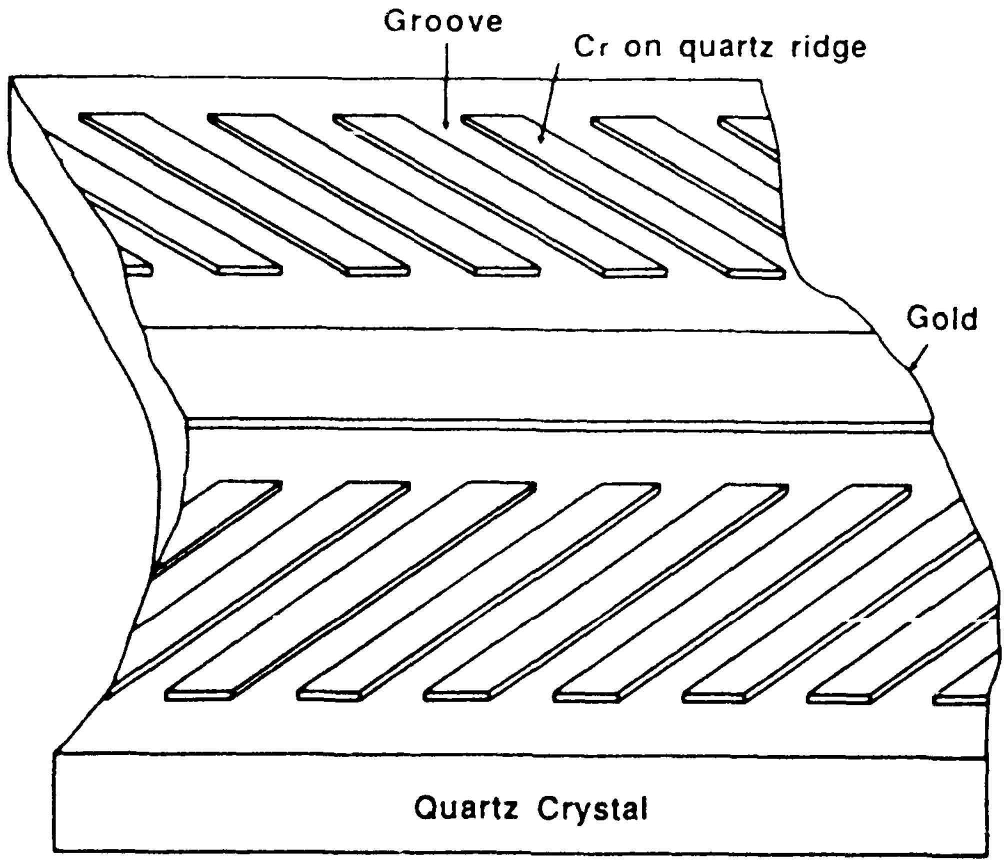


Fig 8 Section of RAC structure with grooves between Cr-covered ridges, the Cr acting as a mask for amplitude compensating etching, and gold film for phase compensation

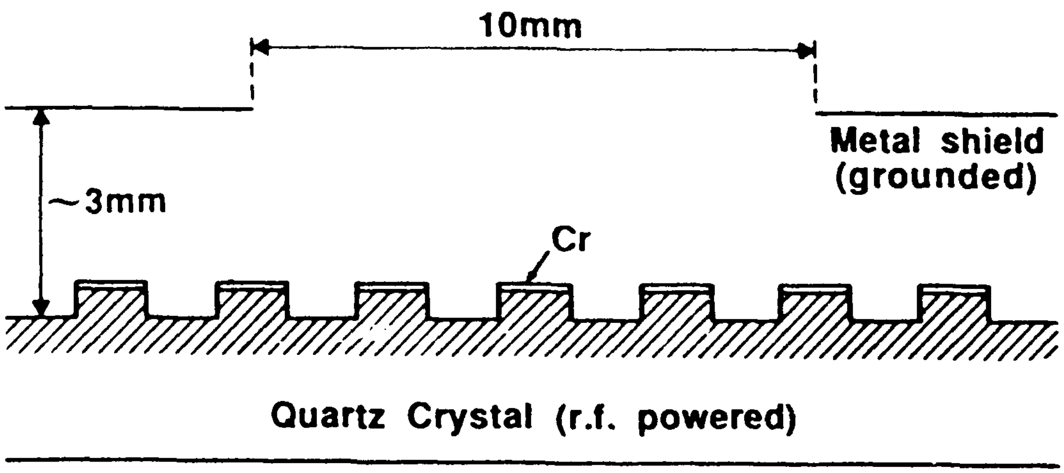


Fig.9 Slit geometry as viewed perpendicularly to the arrays

CFT operates at an IF in the VHF/UHF range. h the output $v_i(O)$ is considered it is seen to correspond to s(τ) filtered by a filter with

MHz processed band be a suitable choice. From extrapolation of experiments. the 10 CFTs required to cover the 2 GHz uplink military SATCOM band would have a total power consumption

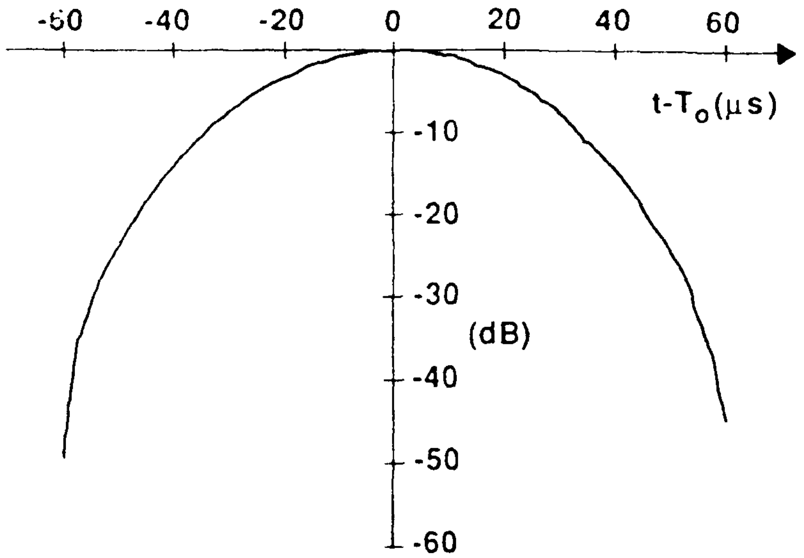


Fig.10 impulse response amplitude versus time as calculated from the measured frequency response Kaiser-Bessel weighted line

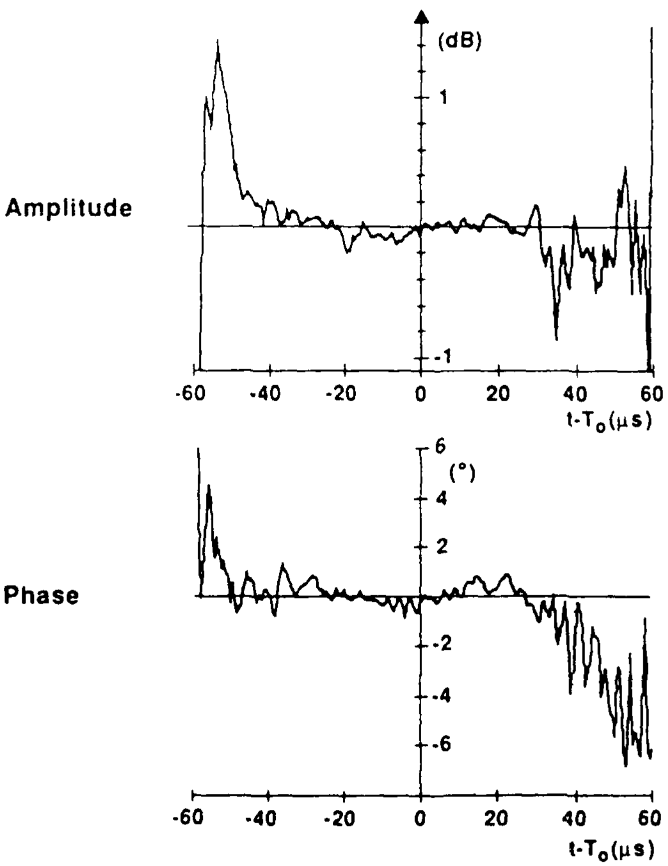


Fig.11 Deviations from desired impulse response amplitude (top) and phase (bottom). Kaiser-Bessel weighted line

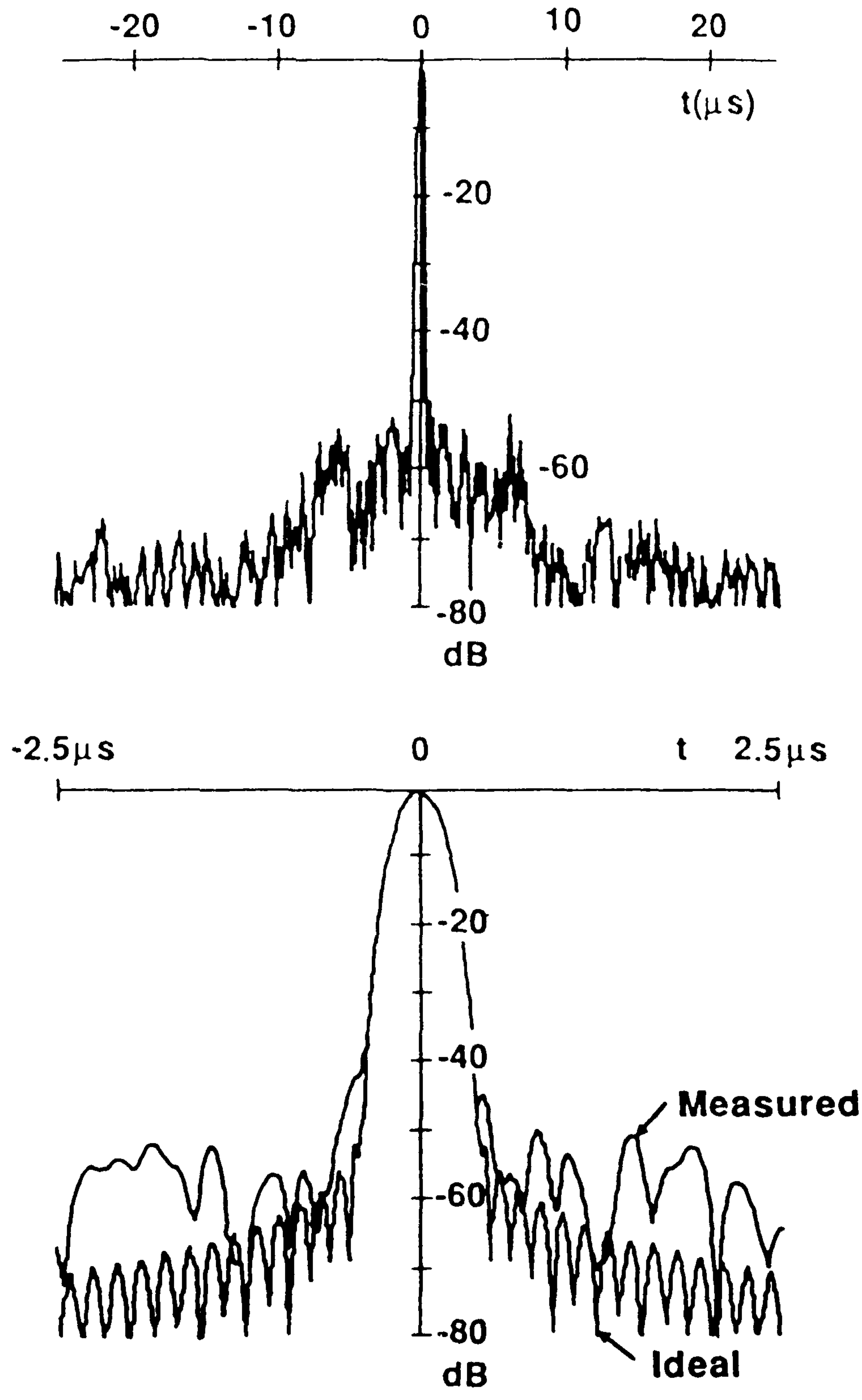


Fig.12 Response from a Kaiser-Bessel weighted line in a simulated pulse compression experiment with the direct feedthrough signal excluded. Top: response over ±25 μs relative to the main pulse. Bottom: response over ±2.5 μs relative to the main pulse for the measured line response ('measured') and the ideal line response ('ideal')

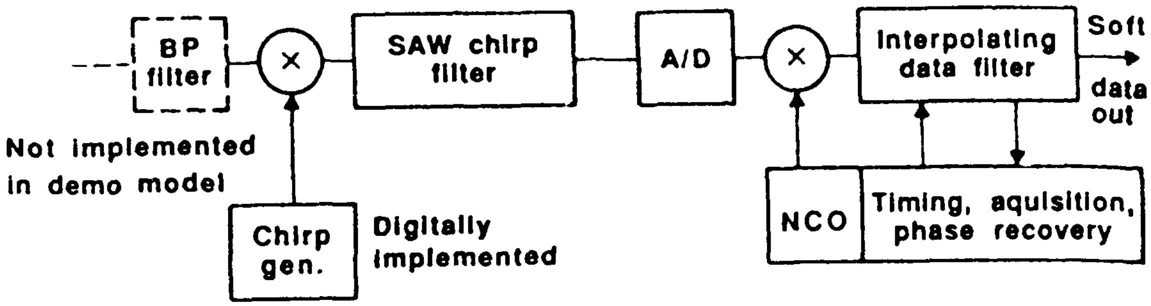


Fig.13 Block diagram of MCD demonstration model built at ELAB-RUNIT

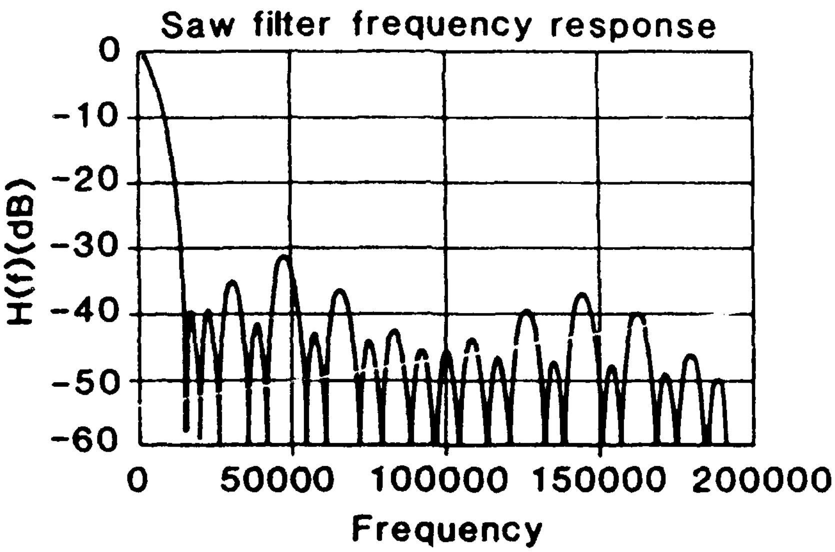


Fig.14 CFT frequency response

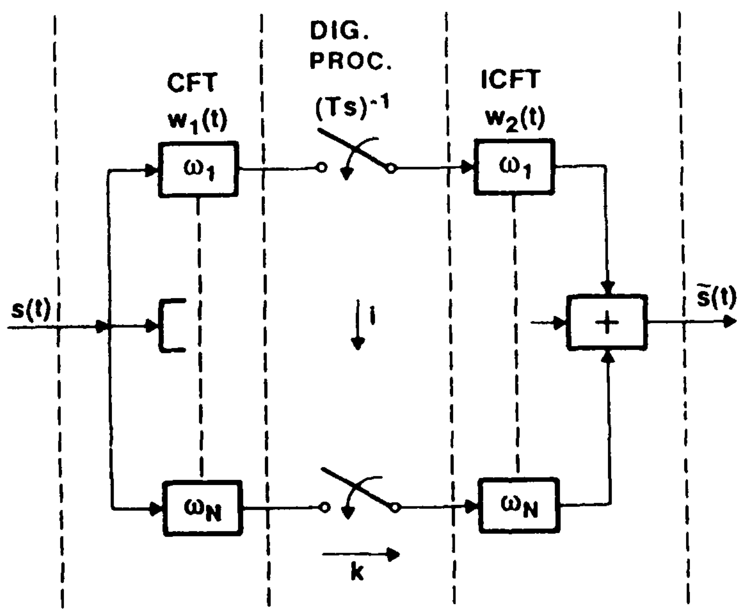


Fig.15 The illustration of equation (21) as a filter bank. The directions of the indices k and i are indicated

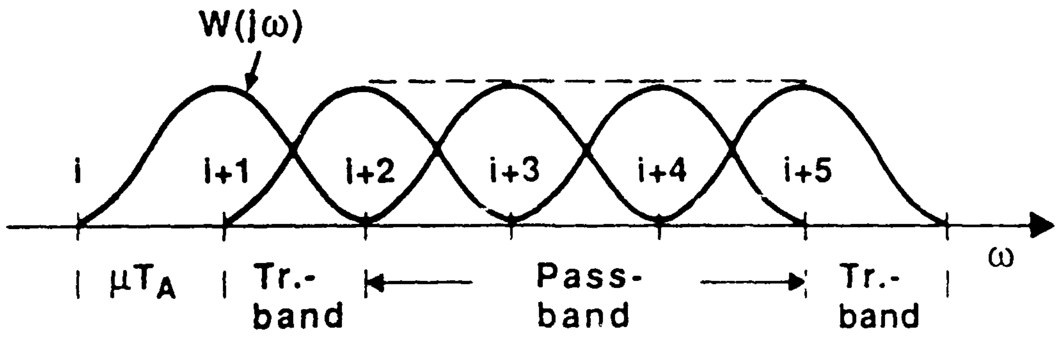
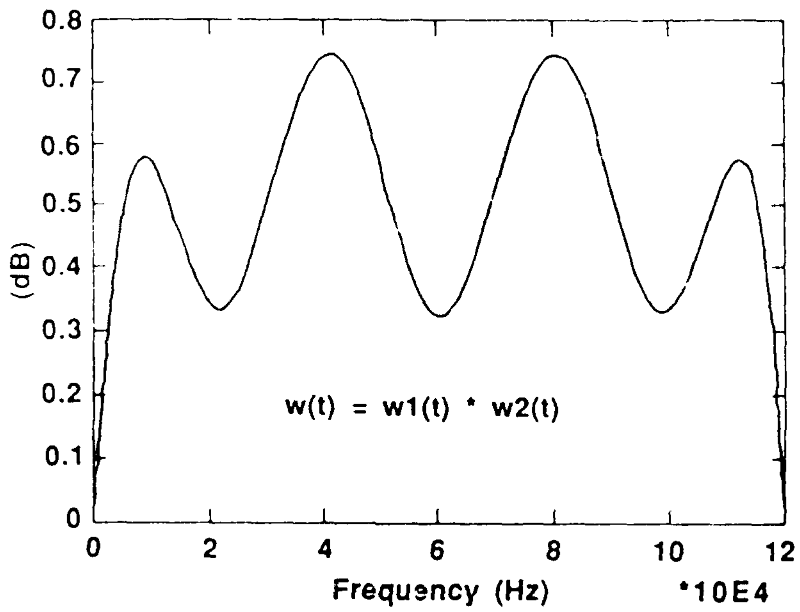


Fig.16 A correct orientation of W(jω) in the filter bank structure



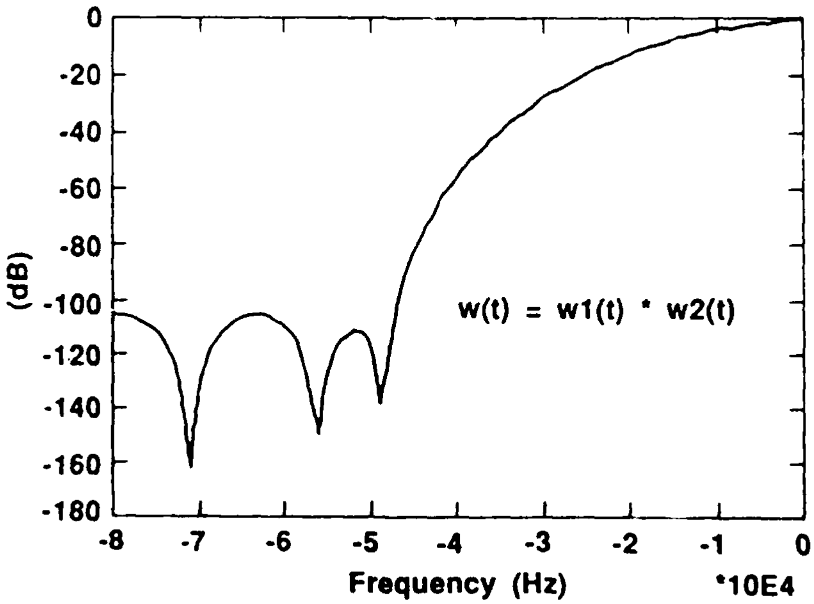


Fig.17 Filter characteristics: (a) passband ripple with four element filters; (b) guardband attenuation with four element filters

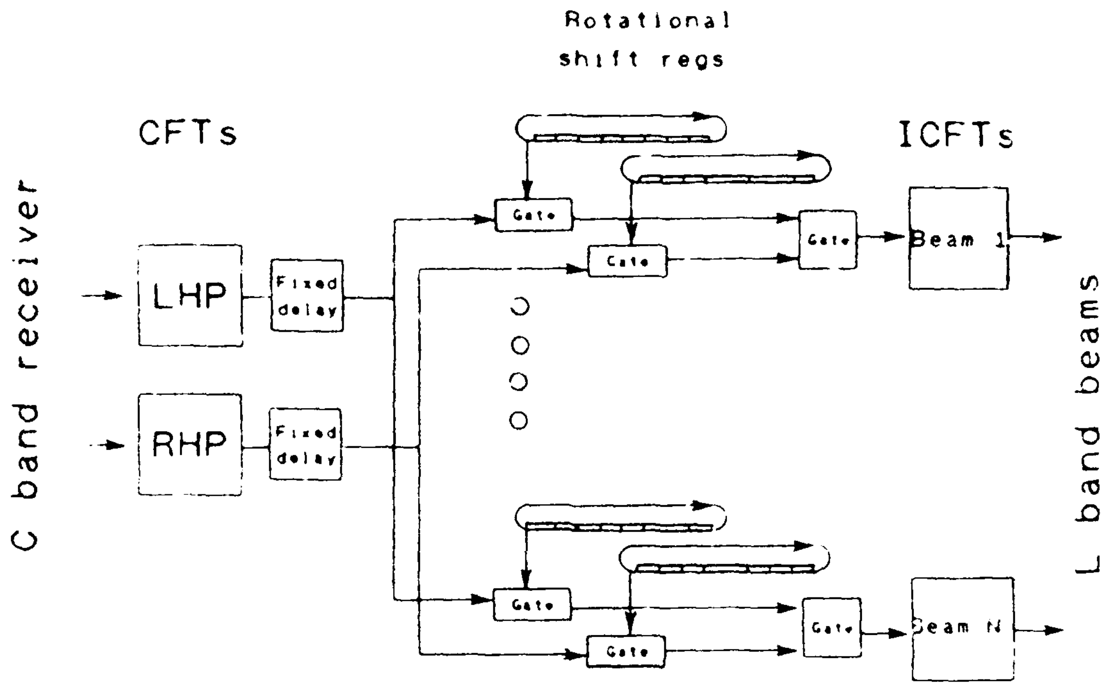


Fig 18 Principle of a simple FROBE for the forward direction. The content of the shift registers is set by telecommand

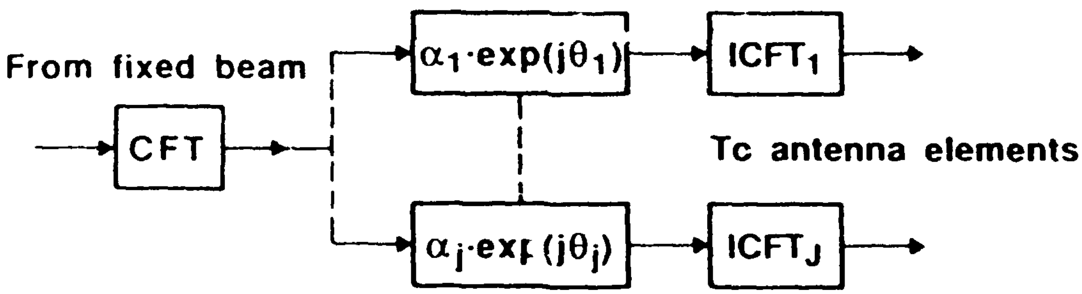


Fig 19 Phased array control principle

8.3 SPACECRAFT PHASED-ARRAY AND MBA ANTENNAS

8.3.1 Introduction

System requirements for airborne, satellite and missile antennas continue to place increasingly severe demands upon antenna technology. In general these requirements push toward the increased capability to control and modify antenna patterns, and away from the use of small antennas with broad radiation patterns. Increased control can imply several levels of added sophistication. At the lowest level it implies mechanical or electronic scanning of an antenna directive pattern, at the next level there are needs to produce precise low sidelobe radiation patterns, and at the highest level of complexity there is the need to actively suppress jammer interference through the use of adaptive control of a full array or an antenna with sidelobe cancellers. In addition to increased control, there is also a trend toward EHF frequencies. At present it appears that the technology at millimeter wave frequencies will be radically different than the current state-of-the-art. System needs at these frequencies are thus seen as a major stimulus to technology (Appendix 8.3.A)

Airborne SATCOM antennas at UHF frequencies were required to have near hemispherical coverage. At L-band there is a need for some limited directivity and possible null steering for anti-jam purposes. At EHF frequencies of 20 and 44 GHz there new requirements for mechanically steered apertures as well as flush mounted arrays. The development of arrays at these frequencies is seen as a major challange to antenna technology. Sidelobe control is not presently an important factor at these frequencies.

Satellite communication antennas have primarily have been lens and reflector systems covering the frequency range mentioned for airborne terminals. At all frequencies the trends are toward electrically larger apertures for increased directivity. Sidelobe control and polarization purity have both been emphasized pushes technology development. Adaptive pattern control has also become a major issue for space segment SATCOM antennas [8.3.3]

8.3.2 Spacecraft Phased-Array and MBA Antennas

Angular or spatial nulling by onboard antennas complements the spread spectrum for AJ protection. There are two basic antenna types for performing this nulling, namely the multiple-beam antenna (MBA), and the phased array (PA). There is sometimes confusion between the two types probably because both contain an array of receive elements, such as horns, that are weighted in some manner to form the desired spatial antenna pattern. The difference is in the plane in which the array elements are located as is illustrated in Fig 8.3.1 and Fig 8.3.2 and discussed below.

For both MBA's and PA's, the output of the N receive elements are weighted in amplitude and phase so as to create, as well as possible, nulls in the direction of jammers while maintaining good gain in the direction of users. In order to chose these complex weights, samples of the received signals from all elements are needed and are processed according to some adaptive nulling algorithm. This topic is the subject of considerable research and is not considered here.

Nulling can be as simple as using a spot beam and relying on low side lobes to supress jammers located there. Here "null" depths are just the sidelobe levels. For full up adaptive nullers, null depths of up to 30dB have been acheived on certain current operational satellite antennas. However, 15 to 20 dB are probably more realistic numbers for all but the most advanced nulling antennas.

There are numerous overviews of MBA's, PA's, and nulling such as in [8.3.1] The debate of the merits of MBA's vs PA's is given in [8.3.2]

8.3.3 Phased Arrays

For the PA, the elements are located in the aperture plane which may be the main aperture itself, as illustrated in Fig 8.3.1, or in the image plane of the main aperture as shown in Fig 8.3.2 (b) This imaging is sometimes used to demagnify the aperture field distribution so that the array elements can be more closely packed. In either case, the received EM field is sampled by the receive elements in one or two dimensions. The individual elements have an antenna pattern that cover the entire field of view (FOV) to be covered by the overall antenna. The samples are of the far field pattern of the transmitted uplink signal. If the samples are spaced sufficiently close to obey the sampling theorem in terms of spatial coordinates, then the Fourier transform, in one or two dimensions, as necessary, would yield a two dimensional pattern that would appear to be the image of the ground terminals and jammers as seen from the satellite. With such an image it is obvious that nulling is easily performed by ignoring the signals coming from the jammer locations. In practice, a two dimensional spatial Fourier transform over the bandwidths required would be diffucult to implement onboard digitally in real time. Instead what is usually done is complex (amplitude and phase) weights are applied to the outputs of the receiving elements in such a manner that the effective receive antenna pattern has nulls in the direction of jammers and as large a gain as possible in the direction of user terminals.

The maximum number of nulls possible with a PA is sometimes given as N-1 where N is the number of elements in the aperture. There is some debate as to whether this limit can actually be acheived in practice.

The number of elements, N, in a PA tends to be large in the hundreds, to give complete and uniform coverage over the field of view. Such a large array is obviously costly in terms of weight and power because of the large number of waveguides, mixers, amplifiers, weights, etc. What is often overlooked is the extra complication of implementing the nulling algorithm which usually requires that N^2 cross correlations be formed, often in hardware. These N^2 correlation values are to be processed according to whatever algorithm is to used such as matrix inversion, gradient methods, etc. Thus, the complexity of determining the weights can be a considerable burden if N is large.

In order to reduce the cost of large values of N, a compromise is usually necessary. The array is "thinned" by removing elements and spacing them further apart than required by the sampling theorem. If the overall aperture diameter is maintained, then the nulling resolution (smallest angular separation between a user and jammer for which the nuller can provide protection) can be mainted. There are several penalties in thinning a PA. One penalty of such thinning is to create what is sometimes called grating lobes within the FOV which actually means that the uniform pattern over the FOV becomes highly irregular. Peaks that now fall in the direction of jammers can be easily nulled by adapting the weights appropriately. However, valleys in the direction of users would be very undesirable because such a user may never be able to communicate. Therefore, designing the pattern for a thinned array becomes very difficult. Another penalty is that with N somewhat smaller, the potential number of nulls, N-1, that can be acheived also becomes smaller. In applications that have a large FOV, such as earth coverage, this is a severe limitation since a large number of smaller jammers can be spread in this area.

8.3.4 Multiple Beam Antennas

For the MBA, the array elements are located at the focal plane of a focussing element such as a lens or, as shown in Fig 8.3.2 a reflector. The field distribution at the focal plane is the Fourier transform of the field at the main aperture. Therefore, the field at the elements is an image of the waves emanating from the earth both user terminals and jammers. Thus, the focussing reflector or lens does beam forming geometrically and is therefore inherently wideband. Much of the phasing required on the PA is done naturally on the MBA.

Each element at the focal plane receives energy from a spot on the earth in a one-to-one mapping from the spot on te earth to the array element, i.e. each spot beam can be associated with a particular array feed element. The beam shape is defined by the diameter of the aperture of the antenna and the pattern of the element itself. Jammers located in beams far removed from a particular feed will be in a far out sidelobe of that feed and, if the sidelobes are sufficiently low, will be effectively nulled.

Jammers located in or near a beam must be dealt with by adaptive nulling. Generally only nearby beams need be used. Typically, groups of 3,7,9 or 19 are used. Three-beam and 7-beam clusters are shown in Fig 8.3.2 where it is assumed that the beams are overlapped at the 3-dB level although this is not necessary for the nulling function. A jammer located in one beam of the cluster will also be received by the other N-1 elements of the cluster through their sidelobes. With these N outputs, adaptive nulling can then be performed to null up to N-1 jammers. Care must be taken to maintain good gain in the direction of the user signals.

Small clusters such as shown in Fig 8.3.3 are the basic units and can be utilized in several ways. In the simples form, a cluster can be used to form a fixed pattern such that the gain in certain areas, such as the Warsaw Pact area, is low. Skynet 4 uses such a format at SHF. In another form, the cluster can be at the center of a mechanically scanned dish so that a composite beam can be mechanically steered to theatre of operation and use the cluster for close in nulling of a few strong jammers. In a more complex arrangement the cluster can be one section of a much larger MBA array-switching is used to select the particular cluster and, therefore, composite beam to be used. Other beams in the MBA can be used simultaneously and independently.

8.3.5 The MBA vs the PA Debate

There is considerable debate over the relative merits of the MBA vs PA approaches with both sides vigorously defending their viewpoints. A signifigant problem in resolving the debate is that even slightly differences in application tend to dramatical tip the balance in one direction so great care must be taken in comparing the two approaches.

In [8.3.2], Mayhan of Lincoln Laboratories discusses the issue at length. An MBA is compared to a thinned PA for the application of providing full earth (18°x18°) coverage with potential jammers located anywhere in this FOV. The MBA and the PA were compared on the basis of three performance criteria:

1) susceptibility to large numbers of interference sources distributed over the visible earth based on a fixed number of adaptive degrees of freedom,

2) trade off between link margin, relative to receiver noise, and nulling resolution;

3) nulling bandwidth constraints imposed by the use of spread spectrum.

The conclusion in [8.3.2] was that for these conditions and performance criteria,the MBA signifigantly out performs the PA. Furthermore,reference [8.3.2] did not consider the extra complexity requried in implementing the adaptive nulling algorithm discussed in Sec. [8.3.3] above which would make the PA appear even worse.
Indefence of PA's they may be attractive for applications that have a somewhat narrower field of view than full earth coverage. It may also have advantage where clever algorithms make the computations for the adaptive nulling more tractable. In summary it looks as though the debate will continue for some time.

8.3.6 References

[8.3.1] L.J.Ricardi, "Adaptive and multibeam antenna systems" Chap 6 in Satellite Communuication Antenna Techonology,R.mittra et al eds.,Elsevier Science Publishers (North-Holland) 1983

[8.3.2] J.T.Mayhan "Area covarage adaptive nulling from geosynchronous satellites: phased arrays versus multiple-beam antennas" IEEE Trans. Antennas Propagation,vol AP-34,pp 410-419,March 1986

[8.3.3] R.J.Mailloux "Basic parameters of antennas" for AGARD Advisory Report No. 232

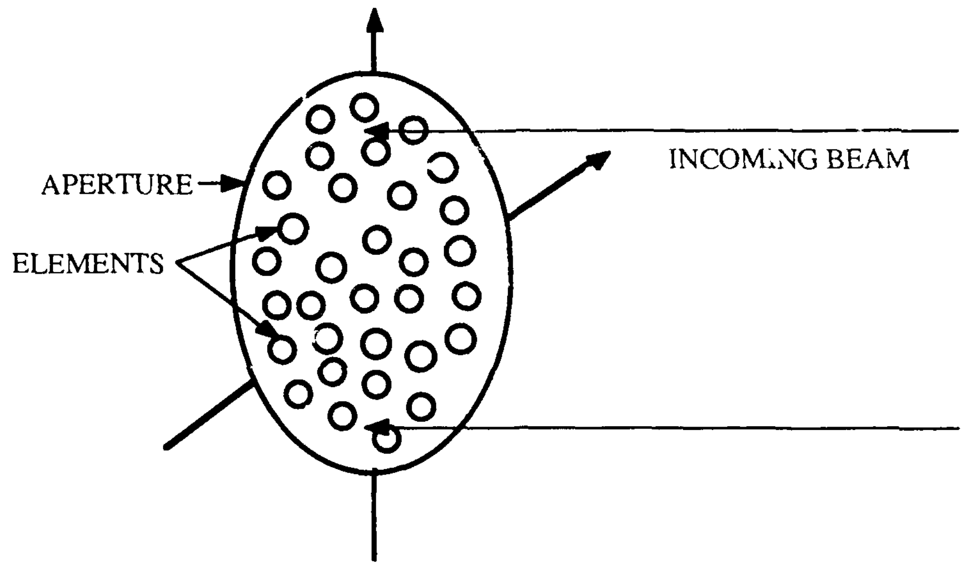


Fig.8.3.1(a) Representation of a PA with elements in the main aperture

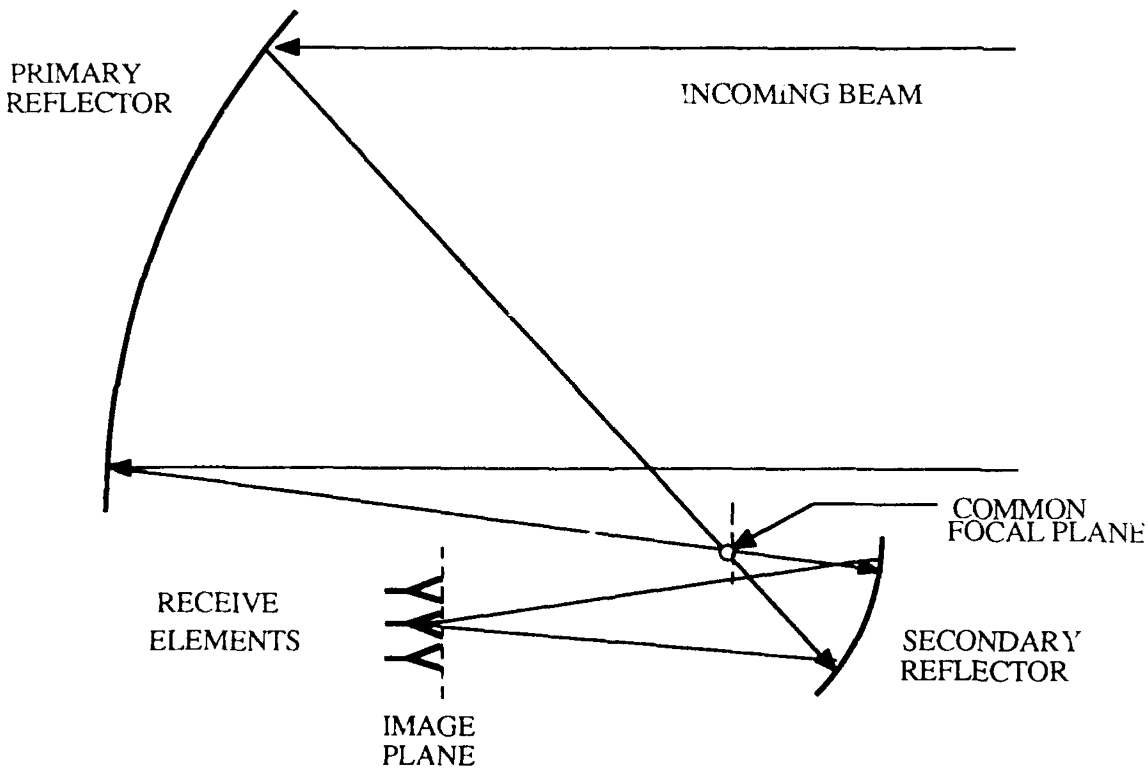


Fig.8.3.1(b) Representation of a PA with elements in a demagnified image plane

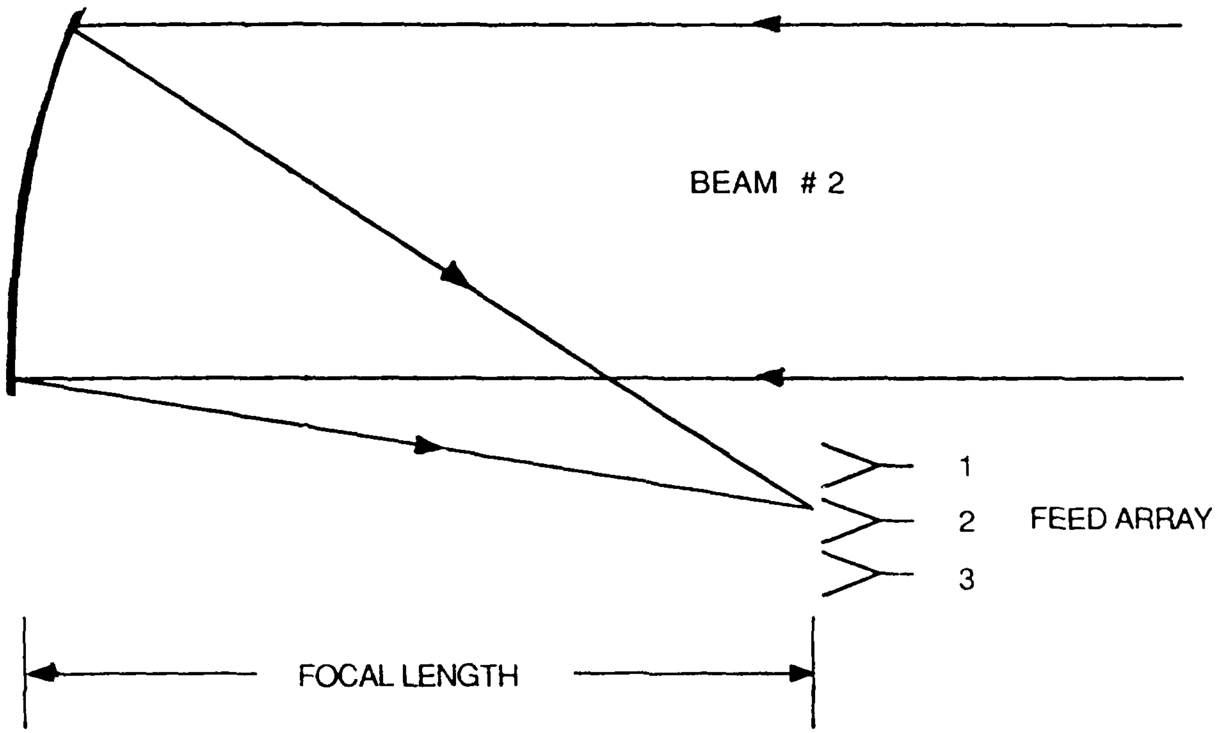


Fig 8 3 2 Illustration of an MBA using an offset paraboloid with feed array at focal plane

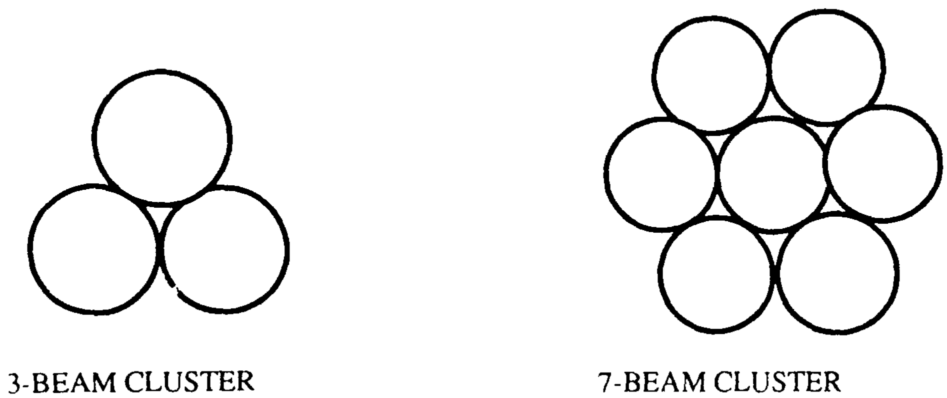


Fig 8 3 3 A representation of the 3-dB contours of the beam patterns at the element array plane

APPENDIX 8.3 A
PHASED ARRAY ARCHITECTURE
FOR MILLIMETER WAVE ACTIVE ARRAYS(*)

This paper discusses general considerations pertaining to the design of active arrays for millimeter wave fabrication. As fundamental examples, microstrip and slot arrays are discussed, and each is shown to impose limitations on pattern coverage for certain choices of substrate or geometry.

1. Active Arrays

1.1 Introduction

There is increasing interest in developing the technology of array antennas of millimeter wavelengths. However, because of the circuit losses at these frequencies, it may be necessary to develop advanced "active" array technology using amplifiers and phase shifters at each element or each subarray instead of merely extending the conventional "passive" (phase shifter) array technology. This is particularly true in the case of large arrays where printed circuit array techniques are required to achieve adequate tolerance and reasonable costs.

In this case the circuits with radiating elements are expected to be collected into groups, or subarrays, on common semiconductor substrate, with microwave power dividers to excite each subarray. Figure 1 illustrates one possible organization of a printed circuit array, and leads to the conclusion that the fabricated array is likely to have a number of printed circuit boards, one for RF power division, one for the control electronics and logic for phase steering, and another for the active elements and subarrays. The design of such multi-layer arrays has to deal with a number of physical problems, like board alignment, thermal expansion and heat removal and especially the maintenance of good electrical connections between boards. The interconnect problem poses fundamental electromagnetic problems and delineates major choices, such as the selection of a printed circuit slot array or a microstrip patch or dipole array, the use of proximity coupling between feed lines and elements (or upper boards) and the selection of what functions are done on what board surfaces. Such basic choices involve very different architectures at the array face, and may seriously limit the performance of the array face as a radiator. Several different printed circuit array configurations will be discussed in this paper.

1.2 Array Architectures

Figure 2 shows four possible configurations of millimeter wave arrays. The upper two use microstrip patch radiators, while the lower two sketches show slot and printed dipole radiating elements. However, the key differences between them are not in the selection of elements, but the choice of semiconductor layer location. The selection of these four configurations is intended only to illustrate a representative cross section of the microwave problems.

Microstrip patch arrays are certainly among the strongest candidates because they are readily fabricated by printed circuit techniques. The upper two figures show the element directly attached to the dielectric substrate thickness is much too thin to support wide band radiation. For these configurations coupling to the elements is either accomplished by plated holes or capacitive coupling. (Fig 2B) or through proximity coupled slots in the lower ground screen (Fig 2A). The semiconductor substrate and active devices are either above (Fig. 2B) or below (Fig.2A) the lower ground screen depending upon thermal costraints and ease of integration with the microwave power distribution network.

(*) Written by R.J Mailloux of RADC, USA

Figure 2D shows a configuration with the semiconductor layer at the radiating array face, and attached to a lower substrate of low loss dielectric material. At higher microwave and millimeter wave frequencies the semiconductor layer thickness may be adequate to support broadband performance with the radiating elements and a ground screen attached directly to opposite sides of the semiconductor substrate. In this case Figure 2D pertains, but without the separating dielectric layer.

Finally, Figure 2C shows one possible slot array configuration with slots proximity coupled to some microwave feed network. An alternative configuration might use probe coupled slots using a microstrip feed attached to the upper slotted ground screen. This orientation has favorable characteristics in that slot arrays are wider band than microstrip patch arrays, and the configuration leaves substantial space on the top surface to insert phase shift and amplifier devices in a planar orientation. One of the key questions concerning this structure is whether it is necessary to create cavities to separate each slot, or whether the circuit of 2C will have good radiating properties without cavity separators.

2. Electromagnetic Studies of Millimeter Wave Arrays.

2.1 Research Activities Stimulated

The above discussion explains why there is an interest in investigating a growing body of array architectures to simplify the EHF array manufacturing processes. Certain configurations have distinct advantages, but it is not evident that any one geometry solves all the problems. Of primary importance are the following practical requirements:

- Minimize or eliminate the use of shorting pins, "plated through" holes and via hole feeds.
- Eliminate the need for forming cavities within substrate layers.
- Minimize multiple layer bonding.
- Integrate thin semiconductor substrate and devices into array.
- Develop architectures with favorable excess heat removal

These issues, essentially arising from manufacturing problems, pose questions of a fundamental research nature that need to be addressed for a wide variety of possible configurations

Among these research topics are the following:

- The study and development of wide band proximity coupling for microstrip and slot arrays.
- Question the need for defining cavities behind scanning slot arrays.
- Investigate surface wave and blindness effects for various array architectures.
- Investigate influence of vertical feed wires on possible blindnesses.
- Investigate the results of using dielectric superstrate layers on array element patterns.
- Investigate feed line radiation characteristics.

Detailed studies of these recearch have not ben conducted for any of the geometries of Figure 2. There are, however some basic data available to describe infinite arrays of microstrip and slot elements.

2.2 Microstrip Array and Slot Array Results

Among the possible architectures for incorporating active devices into the array face, one of the most obvious is to extend conventional microstrip patch and dipole technology. Early results indicate that this is feasible although one must be careful about substrate thickness. The use of thin substrates imposes a bandwidth limitation which may be a limiting factor depending upon the requirements. However. Figure 3, due to Pozar and Schaubert [1] shows that array blindness occurs in such arrays with thick, high dielectric substrates. The reference [1] also points out that these can be avoided, or moved out to much wider scan angles, by selecting a thinner substrate and accepting decreased bandwidth. In addition to bandwidth and blindness concerns, this type of array design remains inherently multi-layer, partly because of the restricted available space at the microstrip patch surface.

Figure 4 shows two types of slot arrays that are often considered for use with achive devices. The primary question addressed in the associated research studies was whether there is always a need for forming cavity separators behind the slots of the array (Case B) or whether it might be possible to omit the separators (Case A). Clearly it would be desirable to avoid this tedious manufacturing process if that were possible. Unfortunately it appears that the cavities are essential to suppress internal modes in the region below the slots. Figures 5a and 5b compare power transmission factors for infinite arrays and show severe blindness for the Case A array with $\epsilon = 9$. The data for the array with cavities (B) and $\epsilon = 2.5$ shows that no such problems exist for that case. However, the power transmission factors for the cavity backed geometry were found only weakly dependent upon the substrate dielectric constant

The conclusion drawn from this study is that some attempt must be made to form cavities behind dielectric loaded slot arrays if they are to be used for scanning. Unresolved research questions pertain to whether one could use microwave circuit techniques (chokes, open circuit lines, etc.) to achieve sufficient decoupling between elements without the use of fabricated cavities

3. Conclusion

The two examples and other geometries referred to in this Appendix show that it is possible to make arrays with high dielectric substrates that exhibit good scan properties. However, the two configurations that have been studied are arrays that are not simply fabricated. Conventional microstrip arrays require feeds that are attached with metallic pins or "plated through" holes, and multiple circuit board construction. Slot arrays appear to require separating cavities that also need precise pin-type fabrication.

Morever, none of the available studies are directly compatible with device integration or deal with thermal constraints. Much work needs to be done to produce array geometries that are inexpensive and simply fabricated, yet which satisfy the need for wide angle, wide band scanning apertures

[1] D.M. Pozar and D. H. Schaubert, "Analysis of an Infinite Array of Rectangular Microstrip Patches with Idealized Probe Feeds", IEEE Trans. AP-32, No. 10 Oct 1984. pp 1101-1107

[2] R.J. Mailloux, "Printed Slot Arrays with Dielectric Substrates", IEEE Symposium on Antennas and Propagation, June 1985

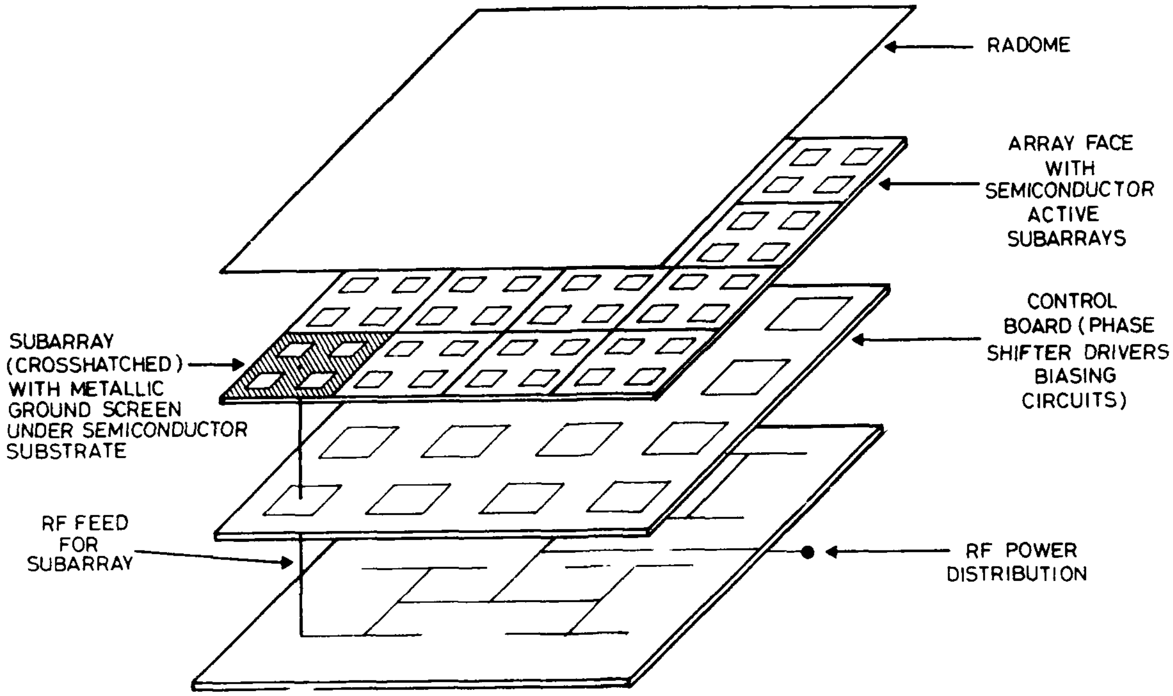


Fig.1 Multilayer construction of printed circuit array

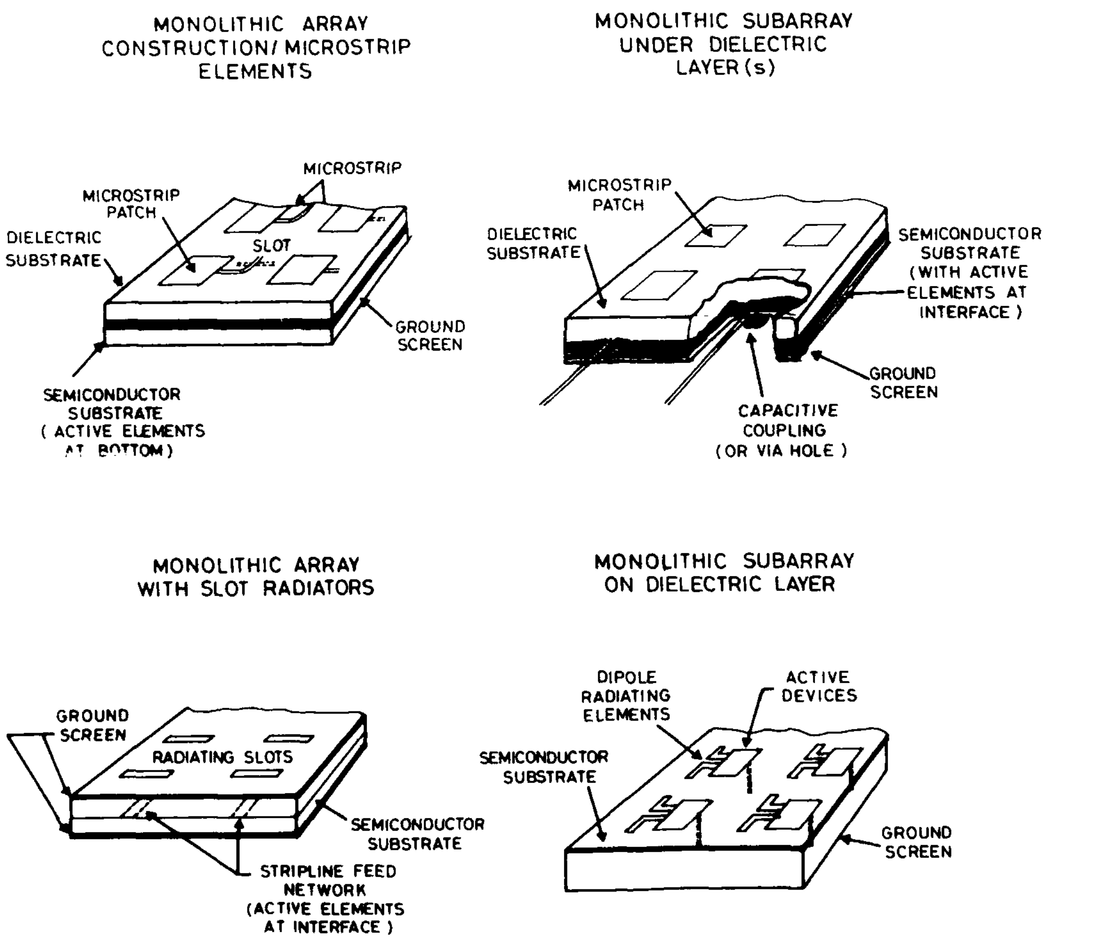


Fig.2 Integration of monolithic MIC technology in ERF arrays

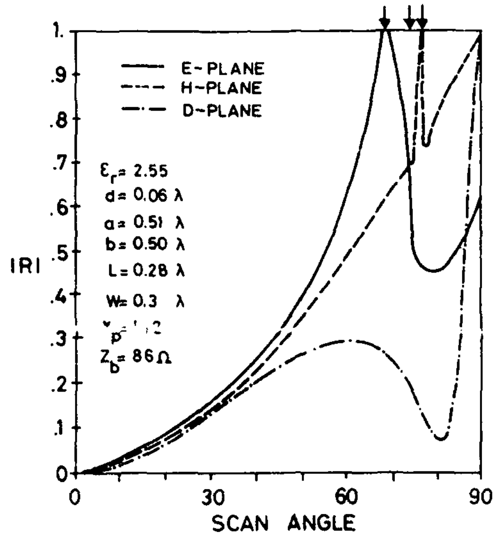


Fig.3 Reflection coefficient for infinite array of microstrip patch radiators (after Pozar and Schaubert, IEEE Trans. AP-32, Det 1984)

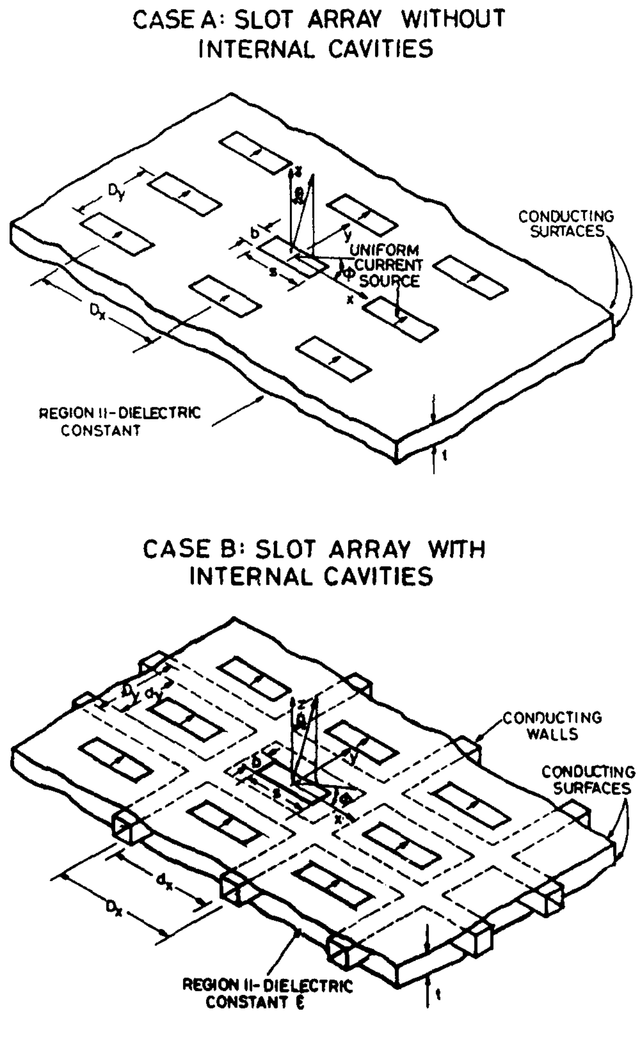


Fig.4 Slot array geometries

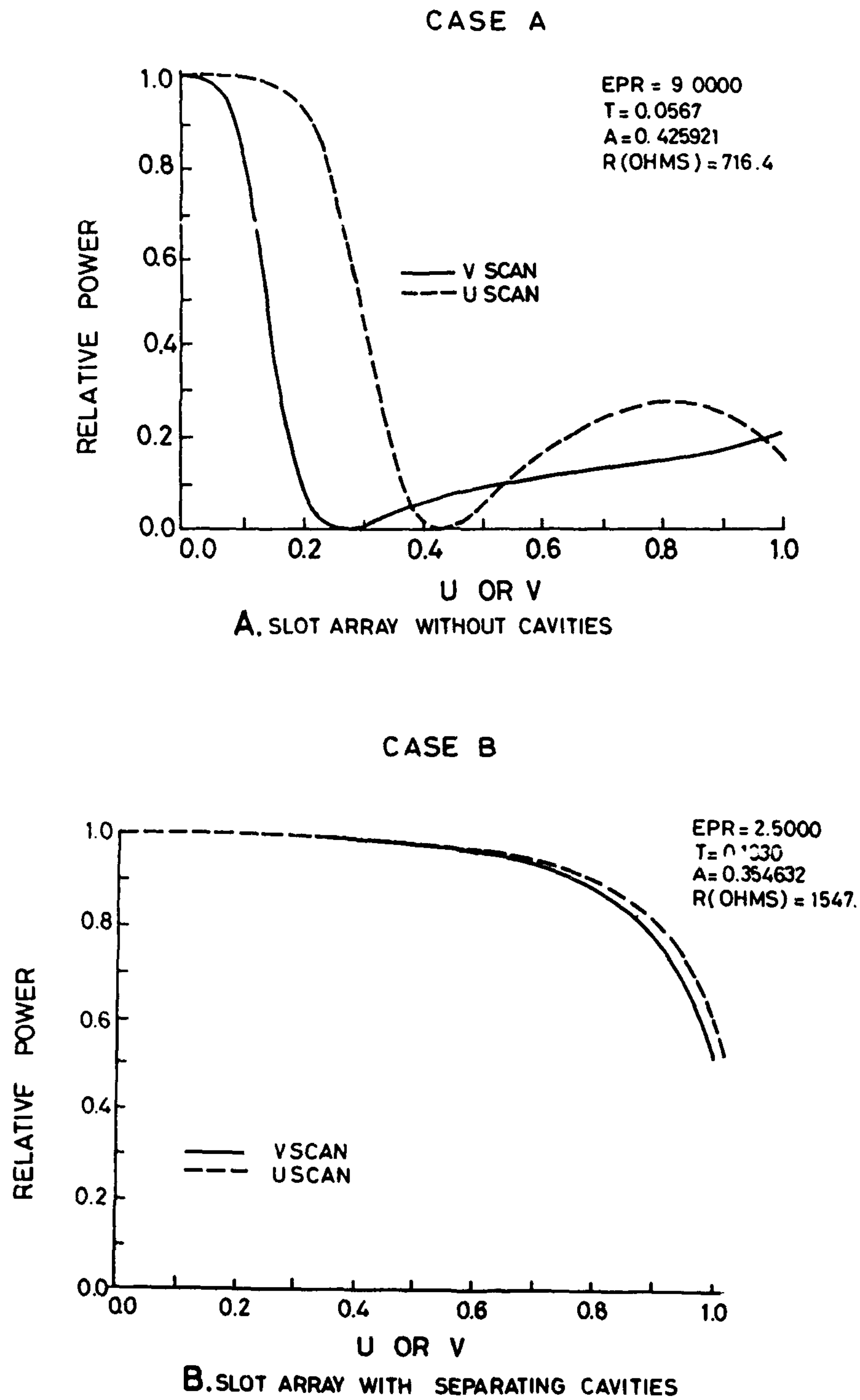


Fig.5 Array transmission factor

8.4 SOLID-STATE SPACEBORNE POWER AMPLIFIERS

8.4.1 Introduction

The basic function of a spaceborne power amplifier is to boost the low-level RF signals provided by the payload low-power section to a level sufficiently high that, when combined with the transmit gain of the Spacecraft antenna, will ensure that the proper effective isotropic radiated power (E.I.R.P.) is sent to ground.

Recent advances in the bipolar and field effect transistor (FET) technologies, particularly GaAs FETs, have had a significant impact upon satellite communications both for earth station and spacecraft applications. Spaceborne power amplifiers, formerly materialized by Travelling Wave Tube Amplifiers (TWTA) exclusively, can now be realized by semi-conductor technologies. Solid State Amplifiers (SSPA) are available to replace TWTAs for low and medium power applications in satellite (all solid state) transponders.

8.4.2 Classification of Power Amplifiers

Power amplifiers can roughly be classified by using the terms output power capability, bandwidth, linearity and efficiency. The requirements and final choice of technology and configuration depends on, above all,

- the antenna type (gain, single/multi feed, reflector or direct radiating antenna),
- the repeater configuration (multiplexing),
- the kind of traffic (telemetry, telephone circuits, television, direct broadcast),
- number of signals (channels) to be transmitted simultaneously,
- signal type (modulation), signal bandwidth and frequency.

8.4.2.1 Output Power

The output power requirements on a spaceborne power amplifier can vary between a few Watts and several hundred Watts.

SSPAs are available today to replace TWTAs for low and medium power applications. Since the late seventies, several hundred SSPAs have been put in orbit on various spacecraft and for various applications as shown in Table 8.4.1. Fig 8.4.1 demonstrates the status and indicates the near future trend (early 1990s) for SSPAs for space applications. A couple of amplifiers in space have been launched for a short term period, only, so as to demonstrate space quality. This is practice in Japan, sometimes.

High power levels mandatorily require high power transistors. The state of the art of microwave transistors is shown in Fig 8.4.3. It reflects above all the performance for commercial parts and, at EHF, the status at R&D. It is supposed that this chart could reflect the technical status of space qualified transistors around the year 2000.

To increase amplifier output power beyond the capability of a single transistor, a commonly used technique is to combine several transistor stages through hybrids. The main drawback of this technique is the large amount of power wasted in the combiners with the respective degradation of efficiency. Fig 8.4.2 shows the combined output power (P_{out}) compared to the power generated in the individual stages as a function of the number of stages and for various insertion losses of the hybrid combiners. Things become even worse if phase differences in the amplifiers and in the combiners are considered. In order not to waste power and efficiency, no more than 4 or 8 stages should be combined.

8.4.2.2 Bandwidth

The bandwidth of an SSPA can be characterized in two different ways:

- static single carrier bandwidth
- and dynamic modulation- or multicarrier bandwidth

The static single carrier bandwidth is solely determined by the microwave characteristic of the amplifier, e.g. transistor parameters, microwave matching networks, power dividers and power combiners. Normaly, it is no problem to achieve a relative bandwidth in the order of 5 to 10 % which is sufficient for space applications.

The dynamic bandwidth of an amplifier depends on the envelope of the microwave signal i.e., in general, the envelope of a multi-carrier signal. This bandwidth must be flat for all frequencies of the envelope. For example, if two carries were applied to an amplifier, the envelope of the composite would offer a frequency component equal to the diference of the carrier frequencies. An amplifier operating in a FDM system has to handle an envelope which contains frequencies starting from DC up to the maximum separation of RF carriers (tens/hundreds of MHz). The amplifier must be able to control its point independent of the envelope spectral properties. With amplifiers operating in a constant DC current mode (purely Class A) there are properties.With amplifiers operating in a constant DC current mode (purely Class A) there are no major problems to achieve large dynamic bandwidths. As Class A amplifiers offer pure efficiency, the high power stages of an SSPA are normally operated Class A/B or Class B. (Class C operation is justifiable for a single carrier, only). The high power stages now draw current according to the envelope of the microwave signal. Consequently, the current in the bias circuits (refer to Figure 8.4.4) will have to follow the amplitude of the RF signal. This is particularly important for the source impedance of the entire bias circuits including that of the power conditioners: The bias circuits must present a very low impedance in order to avoid any drop in the voltage applied to the active device. Such voltage drops, called memory effects, would generate intermodulation distortions (IMDs).

This requirement on the bias supplies is a real technical challenge and cannot be met easily. The use of multiple bypass capacitors, of very large and very small values, connected in parallel in order to bypass the wide range of envelope frequencies, is limited by the fact that high-value capacitors at high frequencies become inductive and form parallel resonant circuits with the lower value capacitors. At such resonant frequencies the intermodulation distortions become very high. Nevertheless, a multicarrier bandwidth in the order of 30 to 100 MHz can be achieved using state of the art transistors in well designed bias circuits.

The bandwidth of SSPAs can be increased by using active bias networks. This technology could be mature before 2000 depending on the need and appropriate funding.

8.4.2.3 Linearity

Device nonlinearities in terms of amplitude and phase cause signal distortions. For example, in multiple-access systems with multi-carrier operation (FDMA) the nonlinearities result in intermodulation distortions. The major causes of intermodulation distortions (IMD) are:

- saturation of the amplifier
- gain reduction and phase shift at high output power levels
- gain and phase distordions at low output power levels due to a gradual and nonlinear change from cutoff to

active mode operation of the transistors.

- frequency dependent memory effects due to non-ideal behaviour of the transistor bias circuits (refer to Section 8.4.2.2)

Real world amplifiers are limiting the signal due to the saturation of the amplifier devices. The IMDs, produced by fractional chipping of the signal, must be taken as unavoidable due to the limited peak power capability of all amplifiers. The amount of intermodulation distortion depends on the signal peak to average power ratio as illustrated in Figure 8.4.5 for the transfer characterictic of an ideal limiting device. The curve represents the theoretical limit.

In fact, the intermodulation distortions of a practical amplifier depends on the magnitude of the device non-linearities (gain and phase). They are practically worse than shown in Fig 8.4.5. The IMDs caused by gain and phase distortions of the SSPA transfer charcteristics at high and low power levels can be minimized by proper linearization of the amplifier. But, special care must be taken to avoid memory effects in the linearizer circuit which again would tend to worsen IMDs at certain envelope frequencies. Figure 8.4.6 shows the typical transfer characteristics of SSPA and TWTA before and after linearization. Figure 8.4.7 represents the corresponding 3rd-order IMDs. When compared to TWTAs, SSPAs offer a much more linear behaviour, and consequently less signal distortions.

8.4.2.4 Efficiency

The DC to RF conversion efficiency of an SSPA is determined by the

- efficiency of the power conditioner
- efficiency of the RF amplifier at the maximum output power
- drop down of efficiency with output power backoff
- and the operating conditions of the SSPA.

The efficiency of todays power conditioners is ranking between 75 and 90 %. The exact value depends on the

- amount of secondary DC power: the higher the power, the better the efficiency. This is due to the power consumption of the conditioners control circuits which is almost independent of the DC power to be delivered.
- number of secondary voltages: the lower the number, the better the efficiency.
- amount of additional control and monitoring circuits, like telemetry/telecommand interfaces and redundancy switching circuits.
- conversion rate of primary to secondary DC voltages: the lower the voltage ratio, the better the efficiency.

The efficiency of the RF amplifier highly depends on the required output power, gain and bandwidth. Table 8.4.2 shows the approximate efficiencies (without power conditioners) at maximum output power which can be expected for different frequency ranges and a gain of about 50 dB. The lower values represent the state of the art, while the higher values tentatively show the numbers expected for the late 1990s.

Frequency	Power/W	Bandwidth /MHz	Max Efficiency /%
L/S-Band	40-60	≈ 50	40-55
C-Band	30-50	50-250	30-45
x-Band	30-40	≈ 300	25-40
Ku-Band	10-35	200-500	20-35
Ka-Band	2-10	≈ 500	15-30

Table 8.4.2 SSPA Performance Summary (at about 50 dB gain)

SSPAs which have to amplify signals with a constant envelope (purely phase mudlated signals) can be operated at a fixed power level close to or at saturation with the associated high efficiency.

SSPAs, installed in a system with amplitude modulation, or amplifying multi-carrier signals, have to be operated in a power backoff mode. Due to the instantaneous envelope of those signals, the operational efficiency becomes the average value of the amplifiers efficiency versus backoff weighted by the amplitude probability of the envelope. Amplifiers exhibit an efficiency decreasing with power backoff. The average efficiency under backoff conditions is therefore always less than the peak efficiency at saturation. The amount of degradation depends on the SSPAs efficiency versus backoff characteristic. The worst degradation happens with amplifiers operated in Class A (constant DC power mode).

Class B amplifiers, where the

- DC power consumption is proportional to the current.
- and the RF output power is proportional to the square of the current.

offer an improved behaviour. Nevertheless, the efficiency drops very rapidly with the squareroute of the power backoff. The drawback of efficiency drop with power backoff inherent in all conventional amplifiers can only be conquered by operating the transistors at voltage saturation even under backoff. In principle there are two ways to do so:

The first one is to make the supply voltage proportional to the envelope of the RF signal. But, this principle has serious shortcomings in terms of

- propagation delays between the modulated supply voltage and the signal envelope
- bandwidth limitations due to detector and modulator circuits
- and, the serious one: State of the art voltage modulators for high power applications are limited to some hundred kilocycles in bandwidth.

The second approach is to control the virtual load impedance of the amplifier by the instantaneous signal envelope over a wide range of power backoff values as shown in principle in Figure 8.4.8. The RF output power and the DC power consumption become proportional to the current and consequently the efficiency remains high. The impedance modulation can be performed by different techniques, like the DOHERTY, the CHIREIX-OUTPHASING and the PAMELA principle. PAMELA has been developed by MBB. It stands for Power Amplifier Module for highly Efficient and highly Linear Applications. The DOHERTY and the CHIREIX-OUTPHASING principle offer shortcomings in terms of linearity, due to piecewise synthesis of the transfer characteristic, and in terms of modulation-and multicarrier bandwidth, due to explicite processing of the signal envelope.

PAMELA applies a technique that overcomes the problems inherent in both other principles. In the PAMELA concept the load variation is performed by three or more individual amplifier branches operating through impedance inverters into a common fixed load, as shown in the principle diagram of Figure 8.4.9 in the case of three amplifiers. All amplifiers are driven by the Steering Network which performes an implicite processing of the input signal envelope. Amplifiers 1 and 3 are driven by signals, whose phases vary in opposite direction as a function of the instantaneous input signal. Amplifier 2 is driven with constant phase. This results in real impedances for the amplifier 2 and in complex impedances for amplifiers 2 and 3.

In order to reduce the average reactive loading to the outer amplifiers which tends to penalize efficiency, both amplifiers are equipped with proper shunt-reactances. From the vector diagram one can see that at phase differences being greater than 90°, negative real impedances would occur causing current flow back

into the amplifiers. Therefore the Steering Network is designed such, that no negative real currents can arise. PAMELA therefore operates as a conventional ampl'fier for very low output power levels. Figure 8.4.10 compares the efficiency for the optimum conventional linear amplifier which is biased under Class B conditions and a PAMELA amplifier. Conventional as well as PAMELA amplifiers are assumed to be free of losses. For ease of comparison, both efficiencies are normalized to 100% at peak power. The conventional amplifier fall notoriously short of performance in this regard.

Design confidence in the PAMELA principle has been established by the development of a breadboard amplifier at L-Band frequencies. The SSPA offers 47 dB gain, more than 40 Watts of output power and a maximum efficiency better than 40 % under single carrier conditions. Loading the amplifier with a bandlimited Gaussian noise signal which is equivalent to an infinite number of individual carriers the following performance has been achieved:

Output Power /W	IMD /dBc	Efficiency /%
15	23	25
20	20	30
25	18	32
30	15	35
35	14	37

The multicarrier performance of the PAMELA breadboard amplifier is superior to that of all other L-Band amplifiers: SSPAs and TWTAs.

8.4.3 Comparison of SSPAs and TWTAs

When compared to TWTAs, SSPAs offer the following advantages:

- better linearity for both phase shift and gain transfer
- less complexity for the power conditioners
- significantly lower weight
- better reliability
- and improved tracking between units.

The disadvantages are:

- less output power capability
- and lower efficiency at maximum output power.

But, in many applications, the disadvantages vanish due to the improved linearity of SSPAs. For example, in FDM systems the power amplifiers have to be operated in an output power backoff mode in order to establish compliance with the intermodulation requirements. To achieve the same level of intermodulation distortions, SSPAs can be operated at 2 to 3 dB less backoff. Therefore, the maximum output power rating of the SSPA can be 2 to 3 dB below that of a TWTA. As the efficiency of an amplifier drops with the power back-off, the TWTA will lose its respective advantage under multi-carrier operation.

SSPAs are likely to replace TWTAs in low to moderate power applications, whereas the area of concentrated high power applications at high microwave frequencies will be left to TWTAs.

8.4.4 Conclusions

During the eighties, a number of SSPAs were placed in space. The quantity was an is still rapidly increasing very fast.

Significant technological progress is going to be achieved with solid state devices. Whatever technology is being applied, amplifier performance can be further upgraded by employing sophisticated amplifier control mechanisms and, last not least, by powerful composite amplifier arrangements.

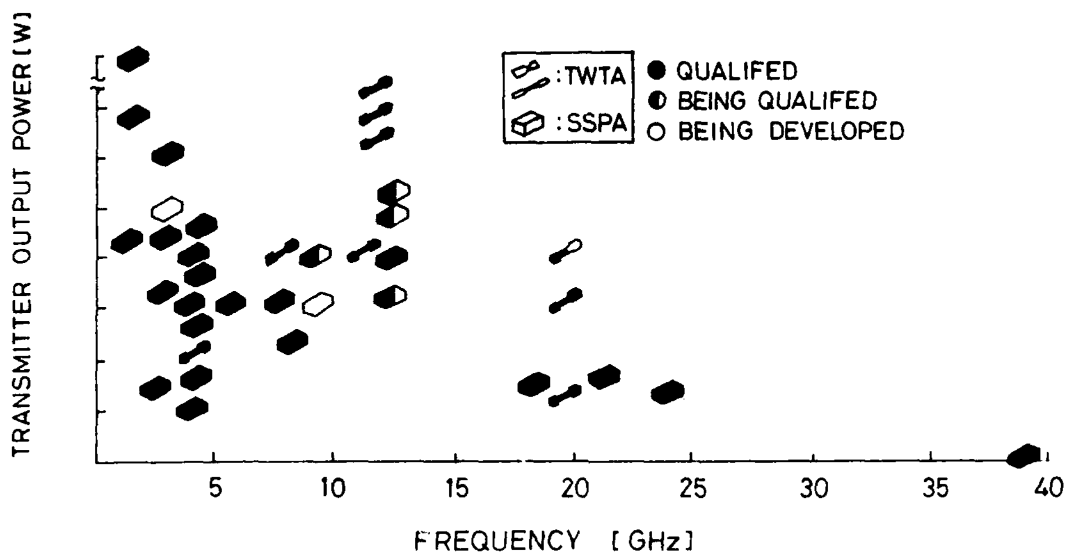


Fig.8.4.1 SSPAs developed for space applications (source: NEC, Japan)

Table 8.4.1
Spaceborne SSPAs

SSPA				
Freq. Band GHz	Power W	Bandwidth MHz	Satellite	Launch
0.25	20 - 65	0.0025 - 0.5	MARISAT	1976
			FLTSATCOM FLTSAT	1978 - 1981
			LEASAT SYNCOM IV	1984 - 1985
			DSCS - III (BLOCK A.B&C)	1982 ...
			SKYNET - 4	1988 ...
1.5	3.0	7.5	ETS.V (EMSS - C. AMES)	1987
	34 / 64	7.5	INTELSAT V (MCS)	1982 - 1984
	75	4.75	MAPECS	1981 - 1984
4	1	?	ETS - IV	1981
	1.8 - 3	72	INTELSAT VI	1989 ...
	5.5	36	TELSTAR - 3	1983 - 1988
			INSAT - 2	1990 ...
	6	180	CS - 3	1988
	8.5	36	ADVANCED SATCOM	1982 - 1986
			ASC	1985 - 1990
			PANAMSAT	1988
			SPACENET	1984 - 1988
	11.5	36	AN.K - E	1990
5	6	22.5	ETS - V	1987
7.5	10	50 - 85	DSCS - III (BLOCK B & C)	?
	32	60	DSCS - III (BLOCK C)	?

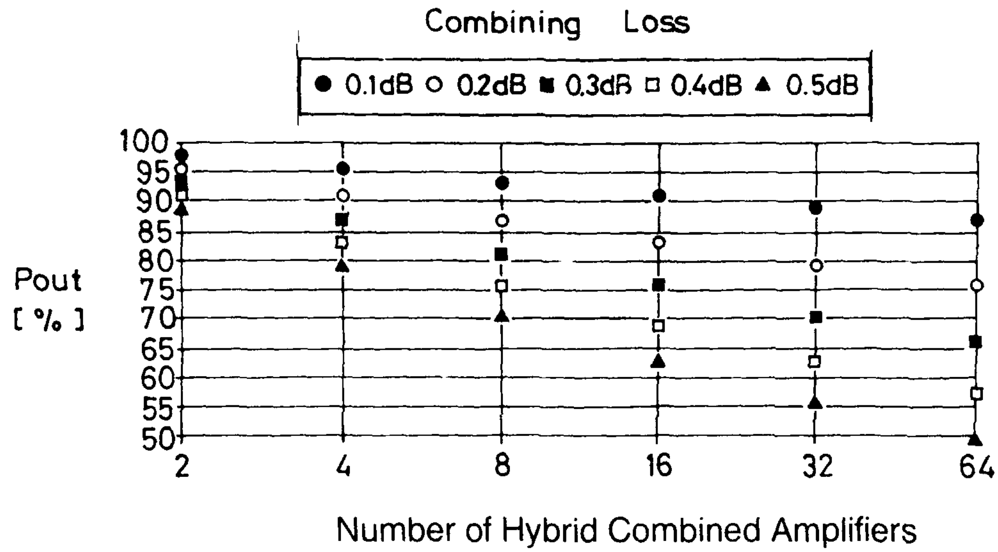


Fig.8.4.2 Combined output power

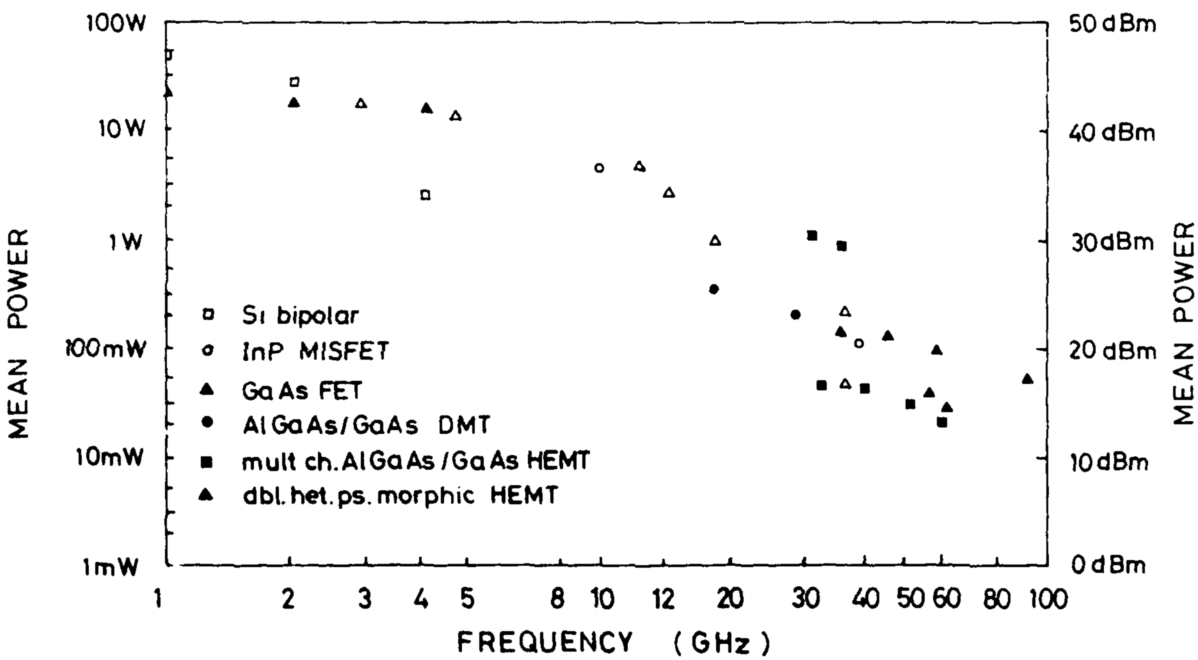


Fig.8.4.3 State of the art of microwave transistors (source: ANT Bosch Telecom, Germany)

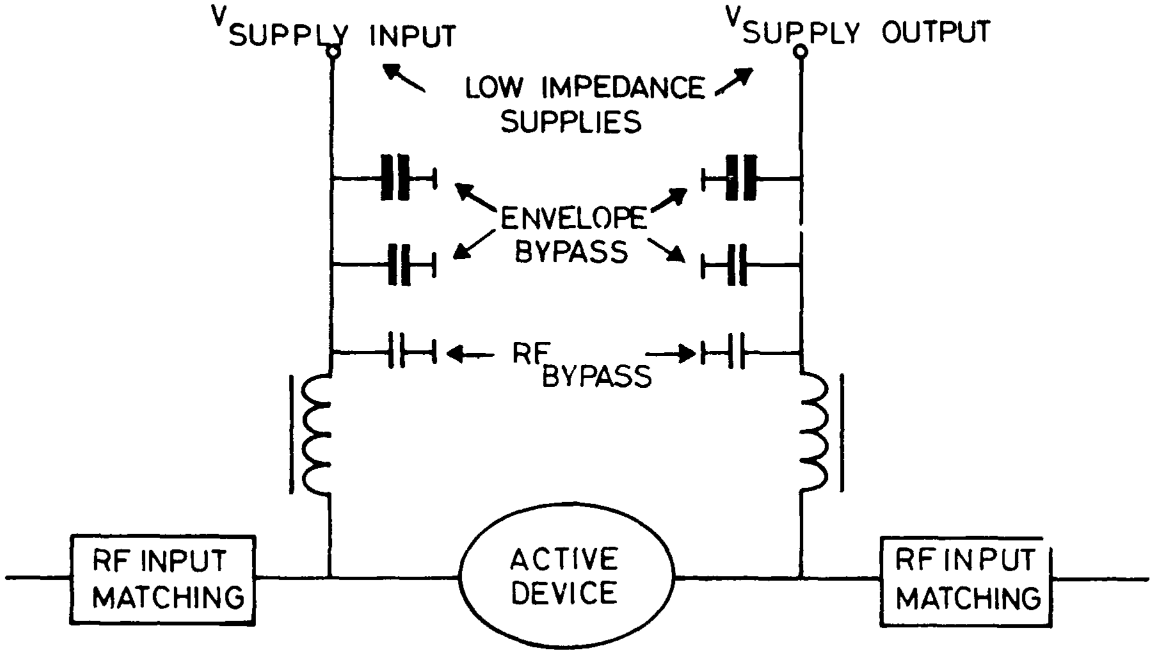


Fig.8.4.4 Typical biasing network for RF transistors

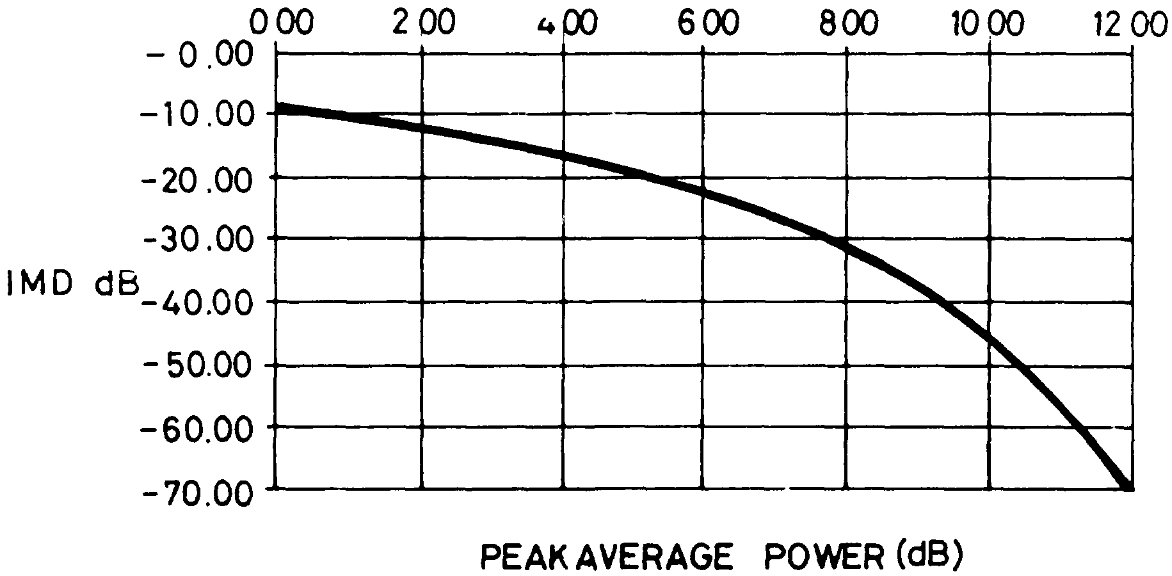


Fig.8.4.5 IMD's expected from fractional clipping

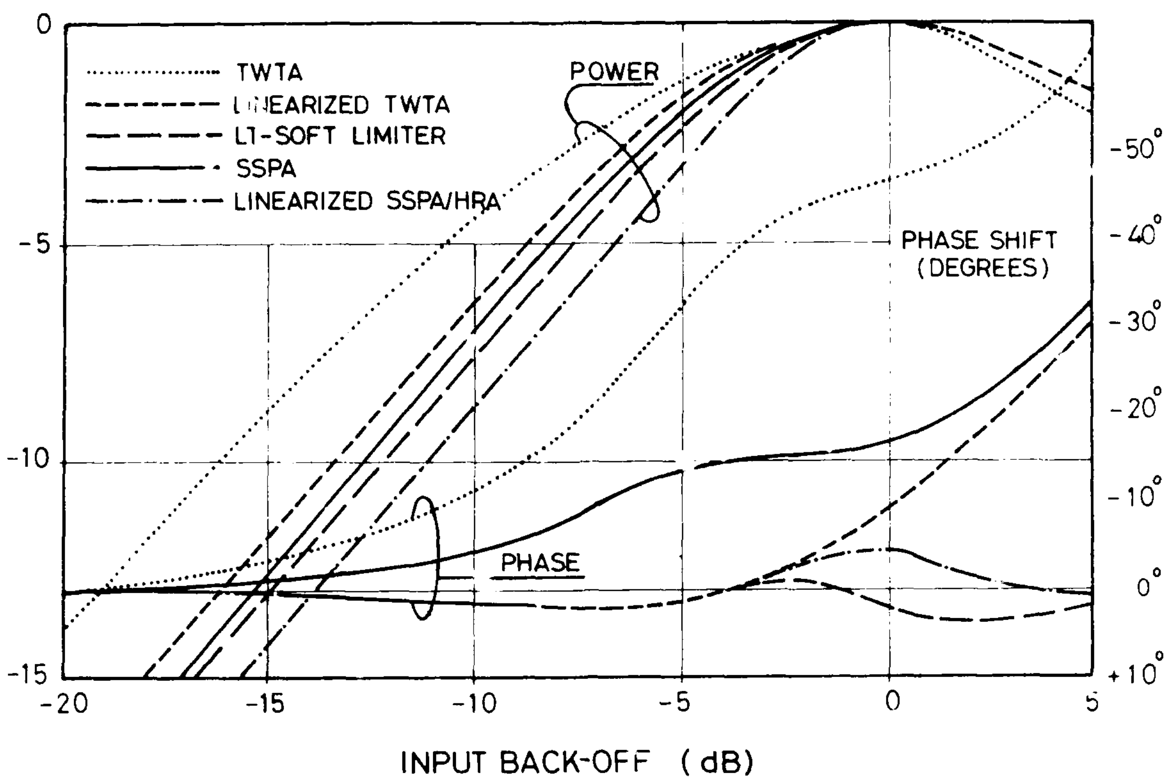


Fig.8.4.6 Transfer characteristics

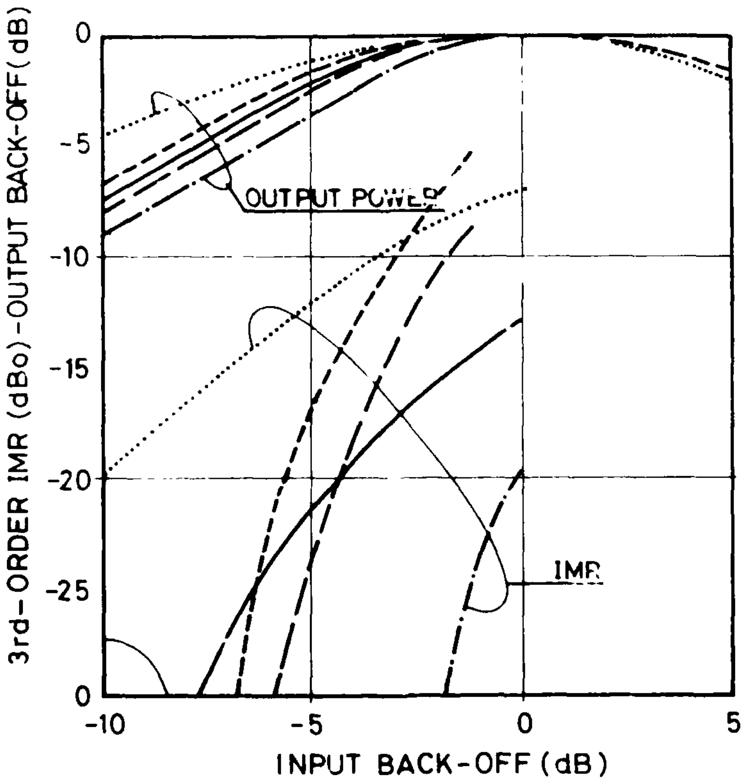


Fig.8.4.7 3rd-order IMD's

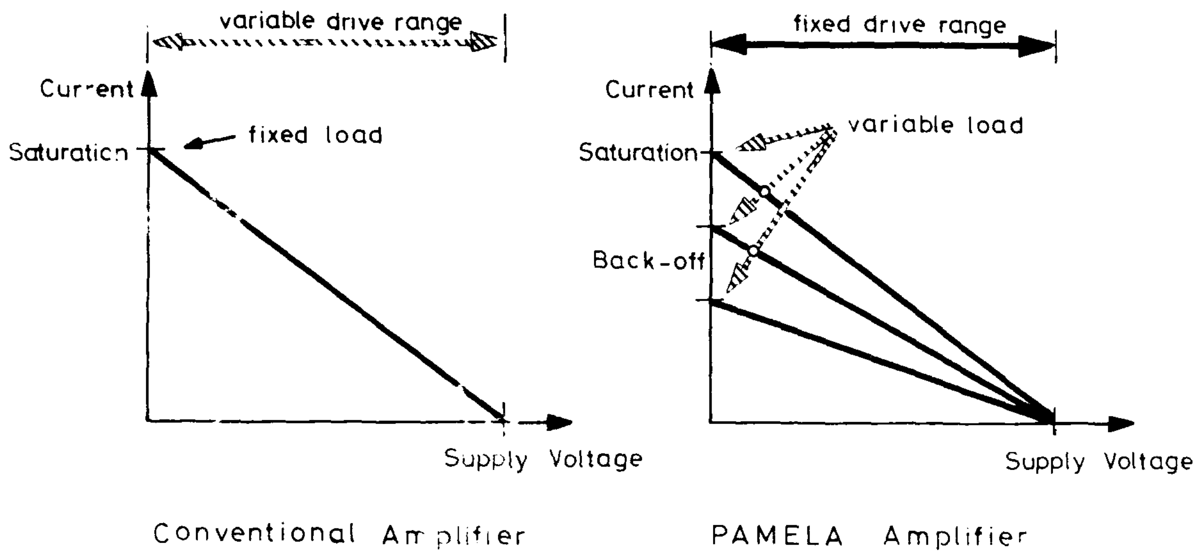


Fig 8.4.8 Output diagrams for conventional and load impedance modulated amplifiers

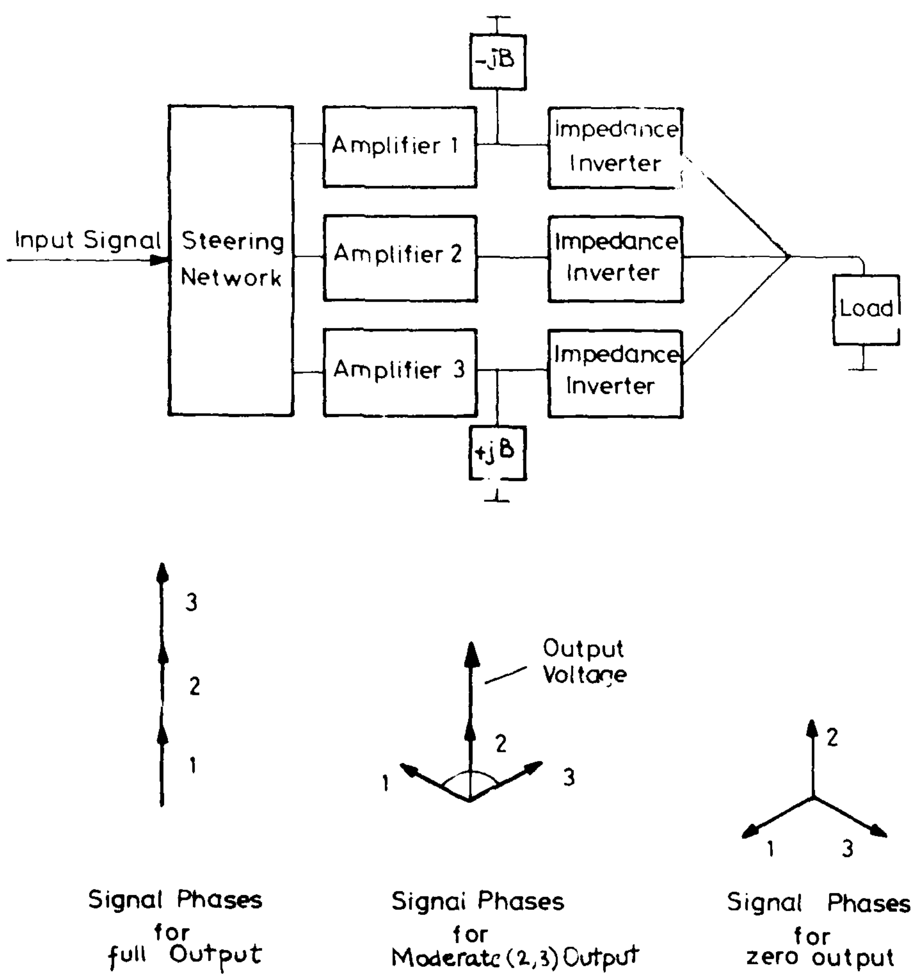


Fig 8.4.9 PAMELA 3-branch system

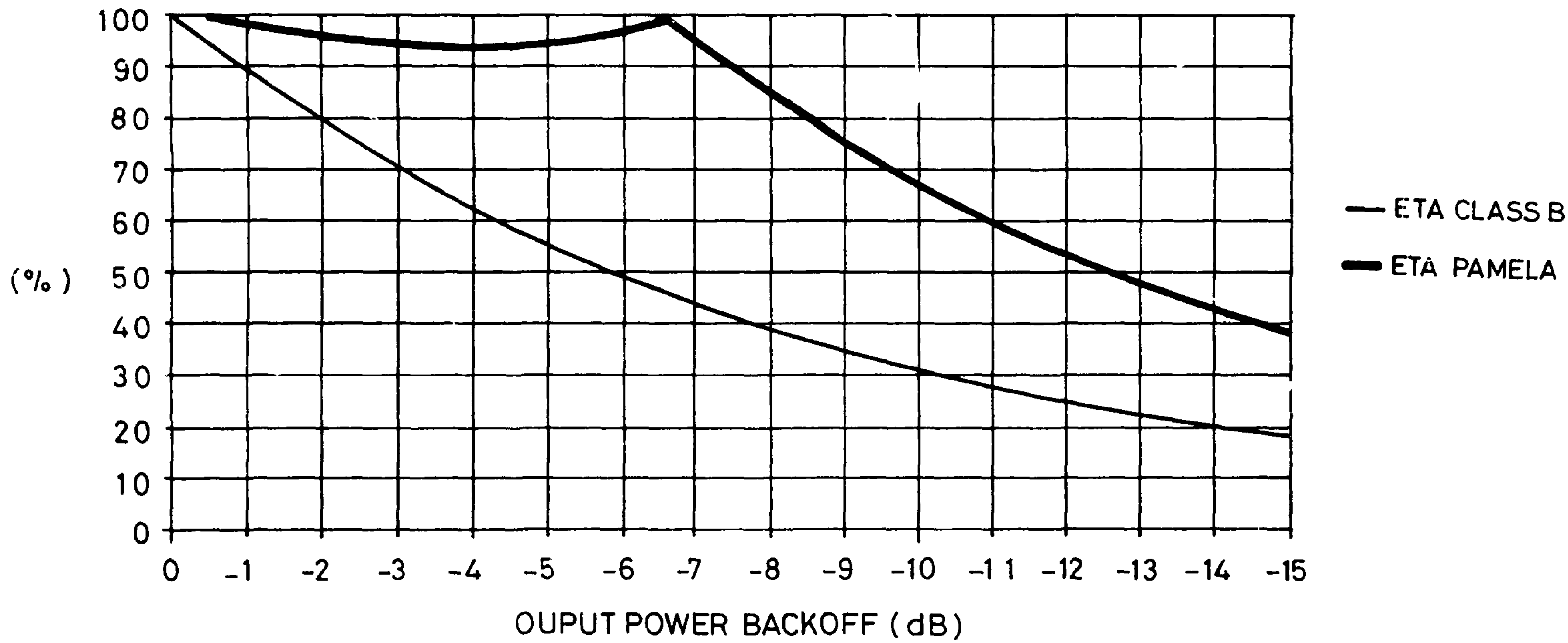


Fig.8.4.10 Efficiency (ETA) of optimum conventional and PAMELA amplifier

8.5 LASER COMMUNICATIONS FOR INTERSATELLITE LINKS

8.5.1 Introduction

Optical communications in space have been discussed ever since the invention of the LASER some 25 years ago. It was recognized early that LASER commonication systems had a potential to transfer data at very high rates. In addition, because of the short wavelength, antenna and beamwidths are much smaller than required for conventional microwave and millimeter wave communication systems. For the next decades, LASER communications are likely to avoid the problem of frequency allocation and congestion, because of the inherent low relative bandwidth and the high spatial discrimination.

Intersatellite communications encompass both Intersatellite Links (ISL) and Interorbit Links (IOL), both illulstrated in Figure 8.5.1 and 8.5.2 respectively. ISL is defined as a link between satellites on the same orbit, usually the geosynchronuous orbit (GEO). IOL refers to satellites on different orbits, usually between GEO and a satellite vehicle on a low orbit (LEO). If not otherwise mentioned ISL is assumed to include IOL, in this report.

The distances over which the LASER links have to work is very long. They are in the order of 72000 km for ISL and up to 44 000 km for IOL. The signal at the receiver is therefore very weak thus leading to very stringent requirements on the LASER communication system in terms of

- optical output power,
- receiver sersitivity,
- beam quality,
- spectral linewidth (as to coherent detection),
- high reliability over long life(> 10 years).

This report will very briefly review the state of the art, and will then highlight the problem areas and promising technological achievements which are likely to be mature for space applications after one to two decades from now.

8.5.2 Rationates for Intersatellite Communications

ISLs can be used to provide extra capacity in satellite networks. Instead of placing a huge satellite in orbit that were to cover all communications needs forecast for the next decade, a cluster concept would offer an economic alternative. If an ISL is used, add-on satellites can be launched to deal with extra traffic and modified connectivities. Another application is the interconnection of similar satellites operating in different regions to provide extended coverage. Also, double-hop link requirements could be resolved by employment of ISL.

The benefits of IOL is being demonstrated over a couple of years through the relay satellite system TDRSS using microwave links. Concepts for follow-on systems are in progress also considering the employment of optical links,see Section 8 5.3.

IOL offers a considerable benefit for the LEO satellite in terms of coverage and accessability. The number of earthstations and the extension of terrestrial networking can be miminized. Storage aboard of the LEO satellite is dispensable in almost all cases.

8.5.3 System Heritage

No experimental or operational LASER experimentation systems have yet been flown in space. However, optical technology has evolved to a point where it is now feasible to consider such systems for space applications.

The ACTS has been designed to carry an optical payload comprised of the main package from the DoD and an add-on package from NASA which together support two different modes of operation: heterodyne and direct detection.

The DoD package consists of a telescope with a beam acquistion and tracking subsysetm, a heterodyne LASER transmitter, and supporting electronics. The telescope has a 20 cm aperture corresponding to a 4 μ rad beamwidth at 0.8 μm wavelength. Beam acquisition is done by conical scanning over the uncertainty region. Tracking is performed by a monopulse technique. The LASER transmitter diode (GaAlAs) outputs 30 mW of optical power near 0.86 μm wavelength. The maximum optical link data rate is 220 Mbps.

The NASA package consists of a direct detection transceiver with supporting electronics and will be linked to the optical subsystem of the DoD package. Each of the three GaAlAs LASER diodes is rated at 70 mW peak power with 50 % duty cycle. With two dioddes combined an output of 120 mW (50% duty cycle) has been obtained. The receiver is designed for direct detection using an avalanche detector (APD) diode with a quantum efficieny of 40 %. The maximum data rate is 220 Mbps. The NASA package has been designed for incorporation onto the next generation TDRS in favour of the IOL. The baseline NASA experiment has been structured for optical links between ACT and two earth stations. Space-to-space links are under discussion between NASA and ESA so as to use one of the planned experimental GEO satellites of ESA, SAT-2 (now TMS) within the PSDE program.

At present the optical ISL package on ACTS has been discarded due to cost reasons.

After the experimental verification of the LES 8/9 microwave ISL, a LASER Intersatellite Transmission Experiment (LITE) has been initiated where a space optical heterodyne system has been designed and developed (App. 8. 5A).

The European Space Agency (ESA) has established a program which also includes an optical ISL payload package, SILEX. It is related to the SAT-2 satellite of the "Payload and Satellite Development and Experimentation" (PSDE) program. Design and development of critical technologies has recently been kicked-off.

Since all the mentioned developments are based on the employment of semiconductor LASERs, the design target is limited to about 100 Mbps.

8.5.4 LASER System Configurations

8.5.4.1 Detection Schemes and Technologies

The receiver takes the incoming optical signal and converts it to an intermediate frequency or a baseband data stream. The performance can be quantified by its sensitivity which depends on the detector material and structure as well as on the chosen receiver architecture.

The traditional materials used for optical detectors, are silikon, germanium, GaInAsP and CMT (cadmium mercury telluride). Figure 8.5.3 shows the normalized spectral response of photo diodes. Silicon can be used in the wavelength range from 0.6 through 1.0 μm and germanium from 1.0 through 1.6 μm. CMT has only been considered for CO_2 LASER systems operating around 10.6 μm. The wavelength range for GaInAsP can be adjusted up to 1.7 μm by altering the detailed chemical structure. A newer compound is GaAlAsSb.

CMT may also be used in the future around 1.3μ using the spin orbit resonance effect. Unfortunately, CMT requires intensive cooling which is heavy in mass and demanding for electrical

power.

The detector device structure also affects the receiver sensitivity. Avalanche photodiodes (APDs) often offer optimum performance although PIN dioes are considered for some wavelengths and receiver types. In particular, the SAM (separate-absorption and multiplication) APD structure currently appears very attractive for future applications.

Finally, there is the receiver architecture itself. The basic LASER communication schemes are based on direct detection Fig 8.5.4 because they are reletively simple in technology but unfortunately, the performance is poor. The modulation is strictly ASK. Much more performant are coherent detection schemes in heterodyne receivers Fig 8.5.4. They outperform the direct detection by 10 to 15 dB, given the same average LASER output power. The modulation type can be FSK, PSK or ASK. The intermediate frequency is usually in the microwave range. Problems are however encountered with alignment of the carrier and the local oscillator, the LASER linewidth and with compensating for the Doppler shift on the incoming signal.

Downconversion to the baseband by mixing the incoming signal with the regenerated optical carrier is refered to as homodyne detection and is about 3 dB superior to the heterodyne receiver Fig 8.5.6 The impact on carrier spectral purity is still more stringent than for heterodyne receivers. LASER diode transmitters cannot meet these requirements, instead solid state LASERS are more suitable.

8.5.4.2 Transmitter Subsystems and Technologies

Four possible technologies are considered for the optical transmitters.

A CO_2LASER transmitter could produce approximately one Watt of optical power at 10.6 μm wavelength. This relative long wavelength leads to a comparatively low antenna gain and the need for a bulky cooled CMT detector. This, together with problems of long term CO_2 molecular breakdown, has rendered this transmitter an unattractive technology.

Much more attractive technologies are based on semiconductors (GaAlAs and InGaAsP) and solid state m+terials (Nd: YAG). Current R&D effort is heavily concentrating on those technologies. The employment of diode LASERs are about to become the short term candidate for spaceborne applications. But physical limits relative to optical power output and spectral purity, may render the solid state device more attractive for mid term applications.

Table 8.5.1 gives a comparative overview on the most promising LASER transmitter technologies and their performance features.

Technology	GaAlAs	InGaAsP	Nd:YAG
Wavelength (μm) (0.946 1.32)	0.8	1 to 1.5	1.06
Optical output power (mW)**)	50	10	1000
Relative reliability	poor	medium	good (due to inherent redundancy)
Efficiency	high	high	medium
Spectral purity			
line width (kHz)	>1000	>100	<100
Rel mass & volume	small	small	large
Beam divergence	large	large	small
External modulator requ'd?	no *)	no *)	yes
Applicable receiver type	heterodyne	heterodyne	homodyne
Possible detection (disragarding ASK)	coher FSK	coher FSK	coher PSK

Table 8.5.1 Comparison of LASER Transmitters for ISL/IOL

LASER diodes of GaAlAs are relatively small in size and have a low power consumption. Also, modulation can readily be achieved by direct modulation of the bias current. Questions, however, remain over reliability and maintaining the output in a single spatial mode at higher powers. Single mode LASERs with an output as given above are commercially available today. Active development is ongoing in the area of higher power LASERs. As much as one Watt may be available in the near future from coherent LASER arrays and/or broad-area devices. But diode LASER arrays are ruled out for coherent detection because their beam quality is not adequate.

GaInAsP diodes have undergone extensive development for the fibre optical repeater market. Initially their longer wavelength and a generally lower power worked against these diodes in comparison with GaAlAs sources. However there is evidence that they can demonstrate higher reliability and can be more suitable for heterodyne receivers.

Nd: YAG LASERs are now attracting ISL applications. A neodymium doped YAG rod is pumped by an external optical source (a flash lamp or a semiconductor LASER array) and can produce one Watt of optical power. The spectral properties are good giving linewidth suitable for coherent detection even with a homodyne architecture. But a number of problems are still associated with the Nd: YAG source: Direct modulation or tuning of the source is not possible.. Some form of electro-optical or acousto-optical external modulator must be used. Also efficiency, especially in the diode pumped case, is relatively poor. Recently a 5 Watt (continuous wave) optical output power has been demonstrated at a French-German joint venture. This has been accomplished with a solid state LASER while maintaining the excellent spectral properties of the Nd: YAG source. Details about the principle are not published yet.

Recent studies agree that the 0.8 μm LASER will be the chosen technology of the nineties. Commercially produced GaAlAs LASER diodes meeting the general requirements of systems in planning, are available today, although further work is needed to fully characterize some aspects of LASER's long term aging characteristics. The CO_2 approach formerly looking promising, is now considered an outdated candidate. Nd: YAG and GaInAsP are becoming increasingly attractive especially when considered in conjunction with heterodyne receivers. They may well become the front runners as IST transmitters in the next century.

8.5.4.3 Pointing Acquisition and Tracking

Due to the extreme narrow optical beam the ISL payload requires a pointing, acquisiton and tracking (PAT) subsystem. This subsystem is still particularly critical for an ISL terminal. There are two ways in which the conflicting requirements of wide pointing angle and high accuracy can be achieved: the indirect and direct schemes.

In the indirect stabilisation approach the wide pointing range is covered by moving a coarse pointing mirror, and a separete fine pointing mirror is to achieve the required accuracy. The fine mirror is to take out any disturbances from the host satellite. The disturbance spectrum is ranging up to 1 kHz. This is a challenging task and is said to be marginally feasible by the fine pointing mirror.

This problem is avoided in the direct stabilisation technique, where the whole PAT subsystem coarsely points to the target. A separete fine pointing assembly is not needed as the high inertia of the PAT subsystem decouples the optics from the high frequency vibration sources on the satellite. The much lower tracking loop bandwidth in the direct scheme is easier to implement and can reject a large portion of the tracking detector induced noise.

*) dc to about 200 Mbps, above that a modulator will be required to keep beam quality

**) At 100 % duty cycle

Work on both types of pointing schemes is continuing. Whilst the direct pointing approach looks increasingly promising. It seems that an in orbit demonstration will be required to settle the argument finally. To substantiate the predicted constraints, ESA has embarked on a Micro-Accellerometer experiment on the OLYMPUS satellite.

Initial acquisition is deemed the most critical process. The rms attitude uncertainty (half cone) may correspond to several hundred optical beamwidths (beamwidth is ranging around 0.3 milli-degrees, antenna aperture around 20 cm). So some form of acquisition protocol separate from normal transmission conditions needs to be established to ensure that the two spacecraft can acquire each other. Beam widening by defocusing of the transmit and receive beam on the respective satellites while running a search and acquisition program on the reciving satellite would eleviate the procedure. The consequence, however, would be that the antenna gains are reduced by a total of around 120 dB. even with a very sensitive acquisition receiver this link loss would cause a link acquisition to take a prohibitively long time.

The use of charge coupled devices (CCD) or charge injection devices (CID) with thousands of pixels offers a partial soultion of this problem as the receiver gain is effectively increased by a factor of the number of pixels. Low-noise devices with as many as 1000 x 1000 pixels have been developed. This approach is currently limited to wavelengths below 1 μm. Secondly, an optical beacon source can be designed to give a much higher beacon source due to its lower duty cycle and coherency requirements. Combining the two methods is likely to result in a feasible approach.

Tracking an optical ISL can be achieved in a manner similar to a microwave system. A small fraction (around 10%) of the communication signal will be used to derive an error signal through a four-quadrant APD or a CCD.

8.5.4.4Modulator Demodulator Architectures

Research on semiconductor LASERs is done worldwide. Such a LASER would be ideal for coherent detection with frequency shift keying (FSK). The frequency modulation is done by direct modulation of the injected current as shown in Fig 8.5.7. It is advantageous to demodulate the frequency at the receiver side by a so called frequency feed-back (FMFB) loop. Fig 8.5.8 shows the schematic of a receiver assembly using FMFB techniques.

Despite the very attractive fetures of diode LASERs such as: high efficiency, long life, and small size and light weight, there are large inherent disadvantages compared to the solid state LASER. Even taking into account further development potential such as high power distributed feedback (DFB) LASERs at 1.3 μm, the power output will hardly be sufficient to fulfil the link budget requirement above 100 Mbps. Diode LASER arrays are ruled out for coherent detection because their beam quality is not adequate

Critical comments can also be made concerning the receiver sensitivity and data rates of coherent diode LASER communication systems. Coherent heterodyne FSK has by theory the same receiver sensitivity as homodyne ASK but remains sub-optimum (by 6 dB) compared to homodyne PSK (phase shift keying).

As heterodyne receivers require a strong local oscillator (LO) LASER, the shot noise generated by the LO also dominates the noise of the subsequent receiver electronics and quantum-limited noise performance is attained. LO intensity fluctuations intrinsic in semiconductor LASERs, can create excess noise in the receiver front end. However, and LO intensity and noise cancellation technique has been developed using the dual-detector receiver concept shown in Figure 8.5.9

The modulation bandwidth of LASER diodes by direct modulation of their currents is below a few hundred Mbps. Modulation beyond this limit is only possible by an external (electro-optical) modulator which basically leads to the systems discussed below.

With GaAlAs LASER diodes, equalisation of the LASER's FM transfer function is required at FSK modulation. This is because unequalized LASER diodes exhibit a large enhancement in low-frequency (< 1 MHz) FM deviation due to thermal tuning effects, resulting in group delay distortions. These effects can be mitigated significantly with passive equalisation of the modulation drive current.

Phase modulation of the LASER diode output is also possible but requires an external modulator or complex injection-locking techniques.

The ultimate objective is the realisation of a transceiver for a LASER ISL using a coherent homodyne scheme and PSK. Such a transceiver would have the highest possible sensitivity and the maximum bandwidth so as to meet any transmission requirement foreseeable for the next decades. Figure 8.5.10 and 8.5.11 show the sshematics of transmitter and receiver assembly, respectively.

Although the hardware complexity in a PSK system is relatively high (Nd: YAG laser, external modulator, and carrier recovery circuitry) DPSK (differential PSK) has considerable relaxed phase stability requirements than PSK and will not require phase locking at high data rates. Open-loop phase stability of diode lasers should be adequate for DPSK modulation at rates beyond 1 Gbps.

The expense for the high performance is the very complex structure of the system. That is why a homodyne ASK is going to be realized in a first step. The requirements on linewidth are relaxed by an order of magnitude. The system is not as complex as for homodyne PSK and it can be realized by components already under development or even available. Figure 8.5.12 and 8.5.13 show the schematics of transmitter and receiver assembly, respectively, for homodyne ASK.

A heterodyne FSK detector is deemed to be possible too, but requires a completely new design of the electro-optical modulator, whereas the modulator for hormodyne ASK and homodyne PSK are essentially the same. Since homodyne ASK receivers have the same sensitivity as heterodyne FSK, the latter is not justified for consideration in future systems.

8.5.5 Critical Items and Conclusion

The outstanding bandwidth properties of optical (laser) Intersatellite Links have raised a number of research and development activities throughout the world. Some benefits can be derived from the achievements of mature optical fibre communications. Above this state of the art, optical ISL requires a higher sensitivity and transmit power as well as a sophisticated PAT subsystem. Current work turns out the most challenging items for future developments. These items are associated with PAT, modem and transmitter (in order of their priority), as discussed above.

Another area requiring demanding optical engineering, is the task of maintaining a low wavefront error. The optical surfaces have to constructed to an accuracy of about one tenth of a wavelength, if performance is not to be degraded by an incoherent wavefront.

A similar problem is encountered in optical heterodying where two laser sources need to be mixed under very tight spatial constraints.

All the principles applied to the design of a microwave ISL and its

hardware are fundamentally applicable to optical ISL. However, the techniques which are available, now and in near future, for an optical implementation are limited. At the current state of maturity optical technology offers a serious alternative to microwave but is not yet clearly superior. The first generation of optical ISL will probably use direct detection and 0.8μm laser diodes. As the problems with more advanced technologies are solved, higher performance systems will become possible. More elegant and performant modulation techniques are about to reach their mutarity for space flight in the time frame of 2000 to 2015 depending on the volume of financial budget made available.

8.5.6 Bibliography

[8.5.1] English, W.:" Study on Intersatellite laser Communication Links," Final Report, ESTEC Contract No. 3555 (1979)

[8.5.2] Leeb, W.R.:" Prospects of laser communications in space," Proc. ESA Workshops on Space laser Applications and Technology, Les Diablerets, 26-30 March 1984 (ESA SP-202, May 1984)

[8.5.3] Kazovsky, L.G.:" Coherent optical receivers: performance analysis and laser linewidth requirements," Optical Engineering, April 1986, Vol. 25, No.4 , pp. 575-579

[8.5.4] Okoshi, T. et.al.:" Computation of bit-error rate of various heterodyne and coherent-type optical communication systems," J. of Optical Communications, 2 (1981), pp. 89-96

[8.5.5] Scholtz, A.L. et.al.:" Receiver Concepts for Data Transmission at 10 Microns," ESA SP-202, pp. 107, 1984

[8.5.6] Philipp, H.K. et.al.:" Costas loop experiments for a 10.6 μm communications receiver," IEEE Trans. on Communications, Vol. Com-31, No. 8, August 1983, pp. 1000-1002

[8.5.7] Zhou, B. et.al.:" Efficient, frequency stable laser-diode-pumped Nd: YAG laser," Optics Letters, Vol. 10, No.2, February 1985, pp. 62-64

[8.5.8] Sun, Y.L. and Byer, R.L.:" Sub-megahertz frequency-stabilized Nd:YAG oscillator," Optics Letters, Vol.7, No.9, September 1982, pp. 408-410

[8.5.9] Ross, M. et.al.:" Space Optical Communications with the Nd:YAG laser," Proc. IEEE, Vol.66, No.3 (1978)

[8.5.10] K. Wallmeroth, "High Power, CW Single-Frequency TEM.., Diode-laser-Pumped Nd:YAG laser" Electronics Letter, 18.8.1988, Vol.24, No:17.

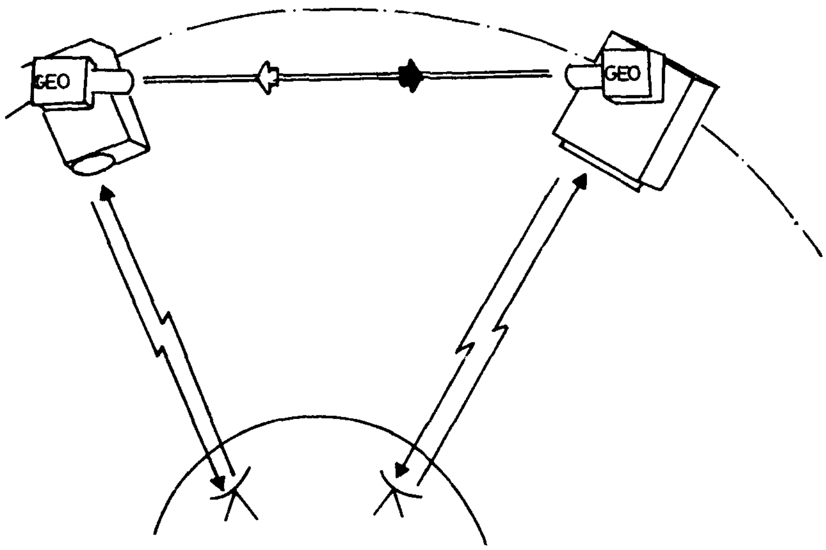


Fig.8.5.1 Schematic of an intersatellite link

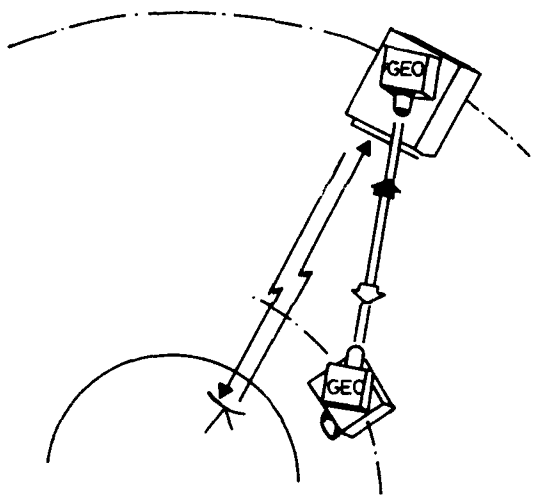


Fig.8.5.2 Schematic of an interorbit link

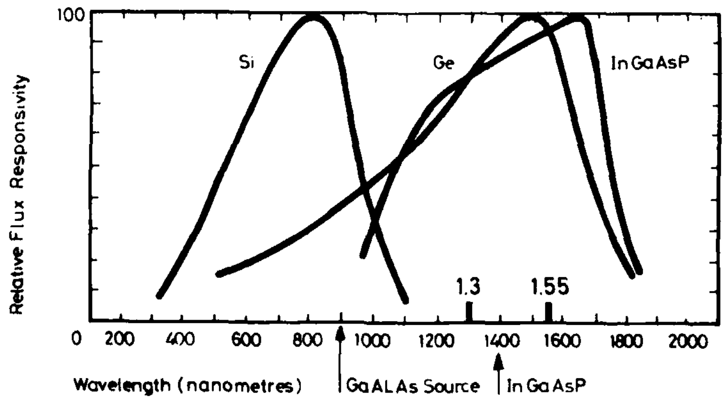


Fig.8.5.3 Spectral response of photo diodes

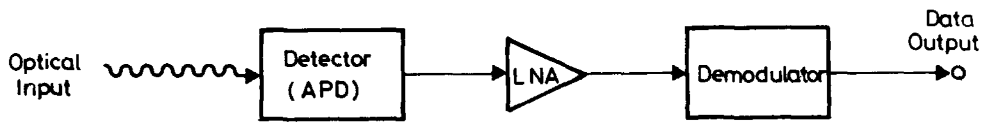


Fig.8.5.4 Schematic of a direct detection receiver

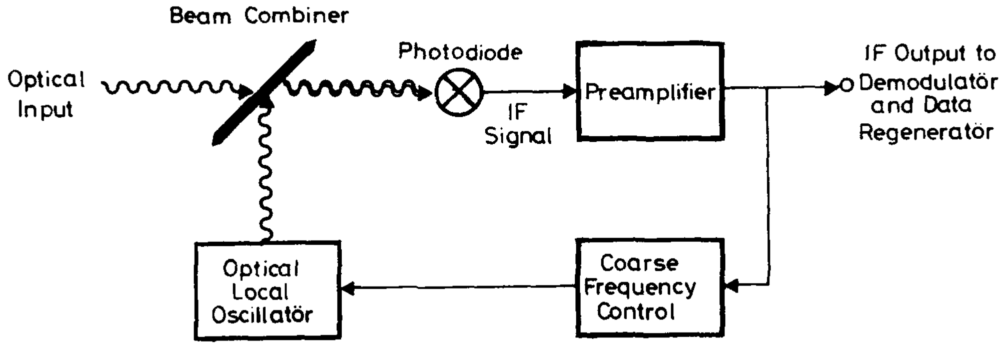


Fig.8.5.5 Schematic of a heterodyne receiver

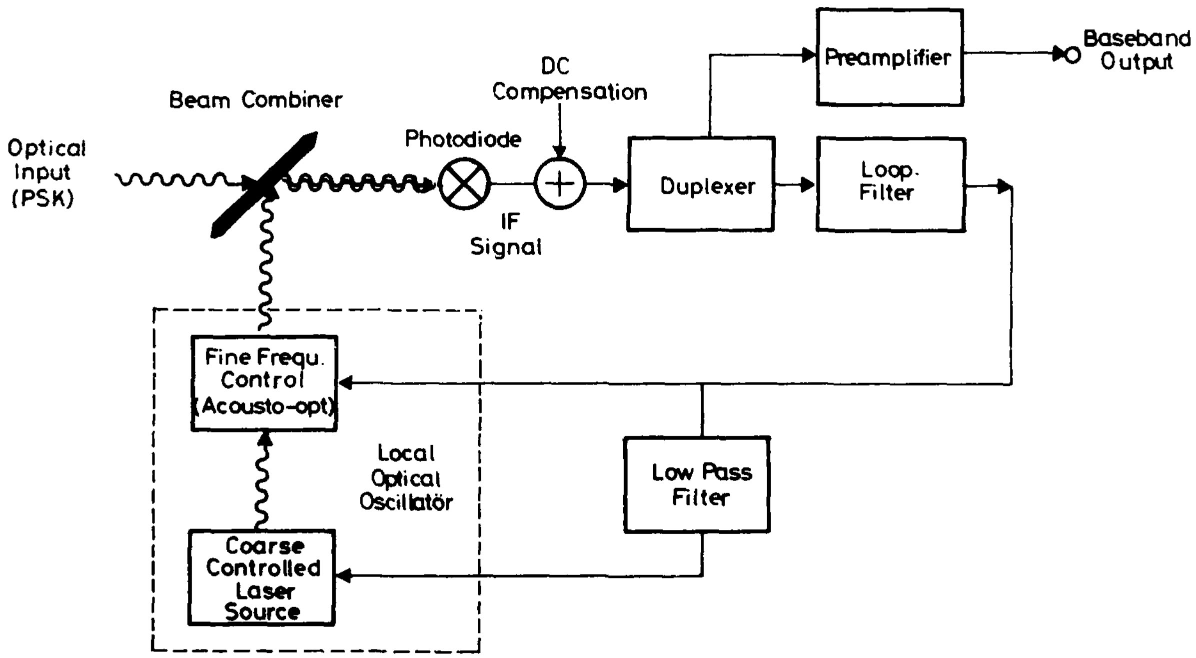


Fig.8.5.6 Schematic of a homodyne receiver

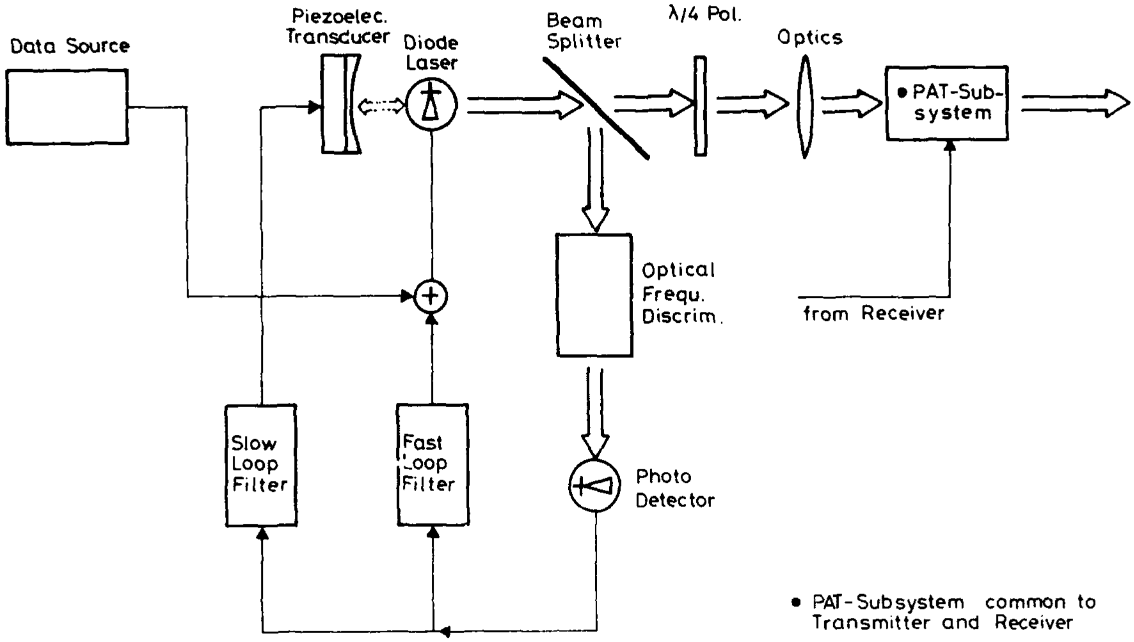


Fig.8.5.7 Diode LASER transmitter assembly (FSK)

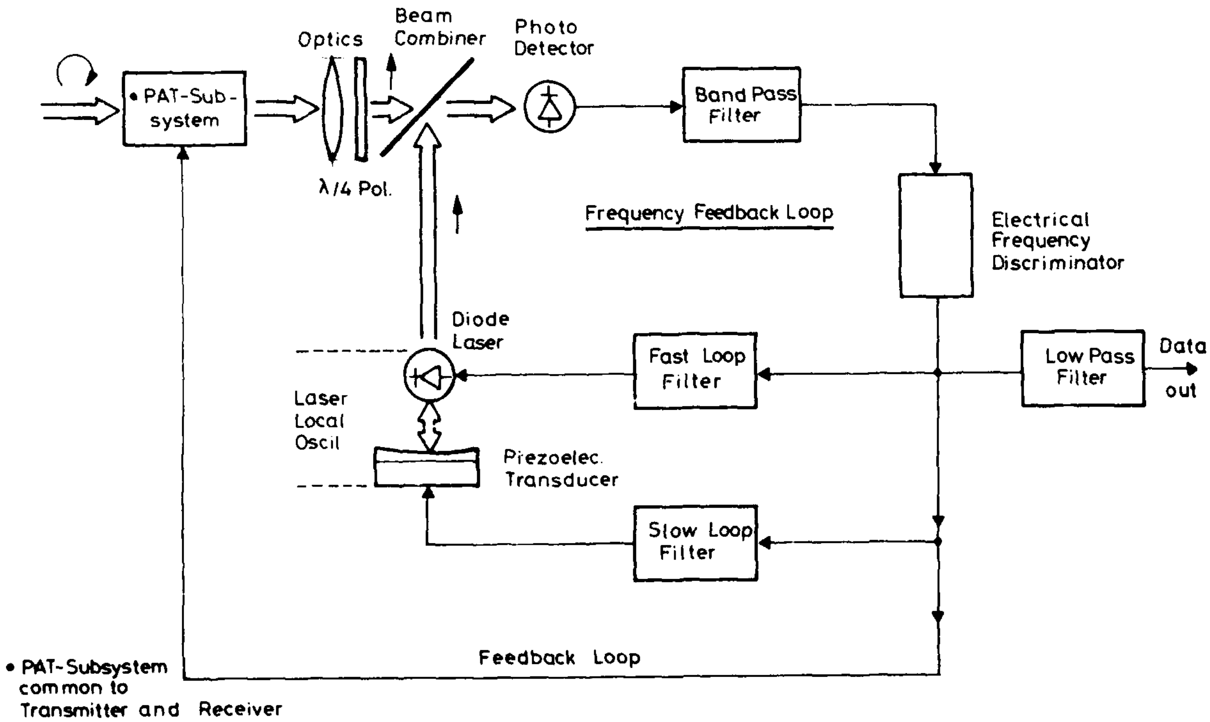


Fig.8.5.8 Diode LASER receiver assembly (FSK)

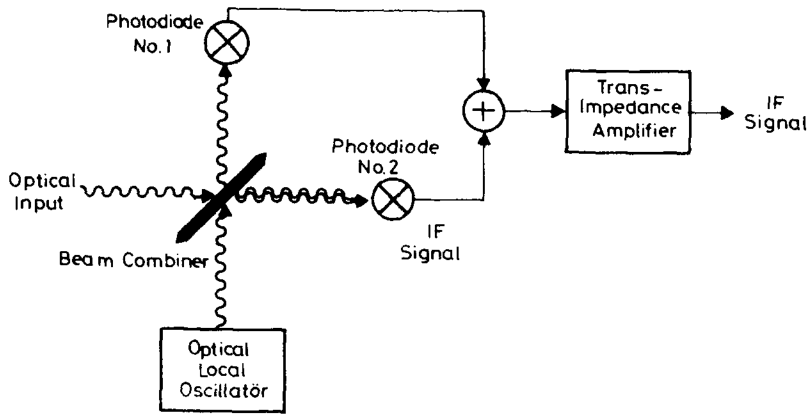


Fig.8.5.9 Concept of LO intensity and shot-noise cancellation technique

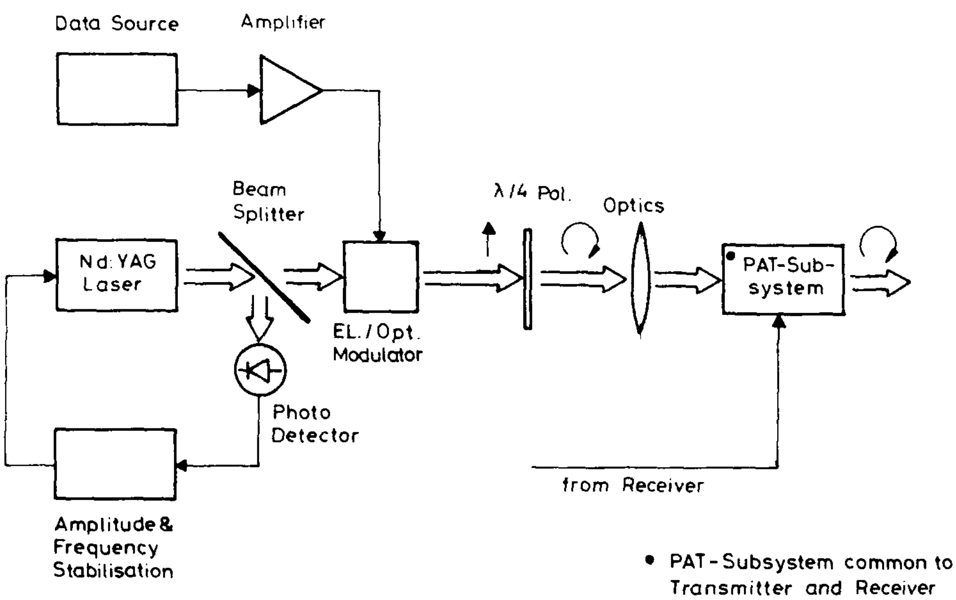


Fig.8.5.10 Nd:YAG LASER transmitter assembly (PSK)

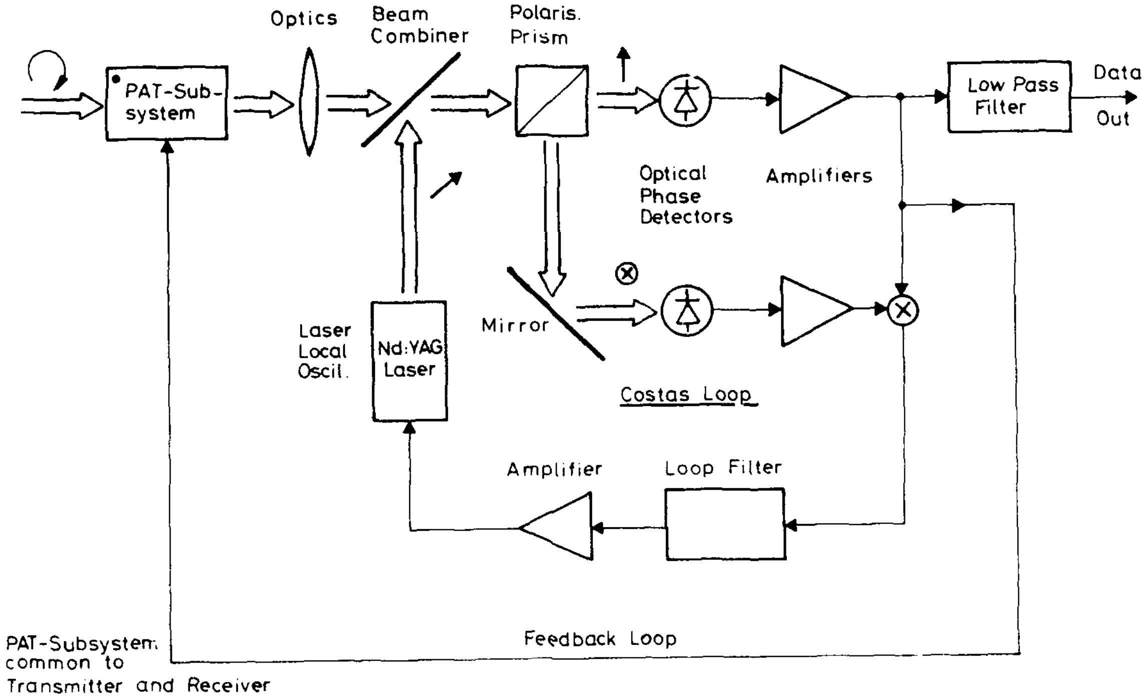


Fig.8.5.11 Nd:YAG LASER receiver assembly (PSK)

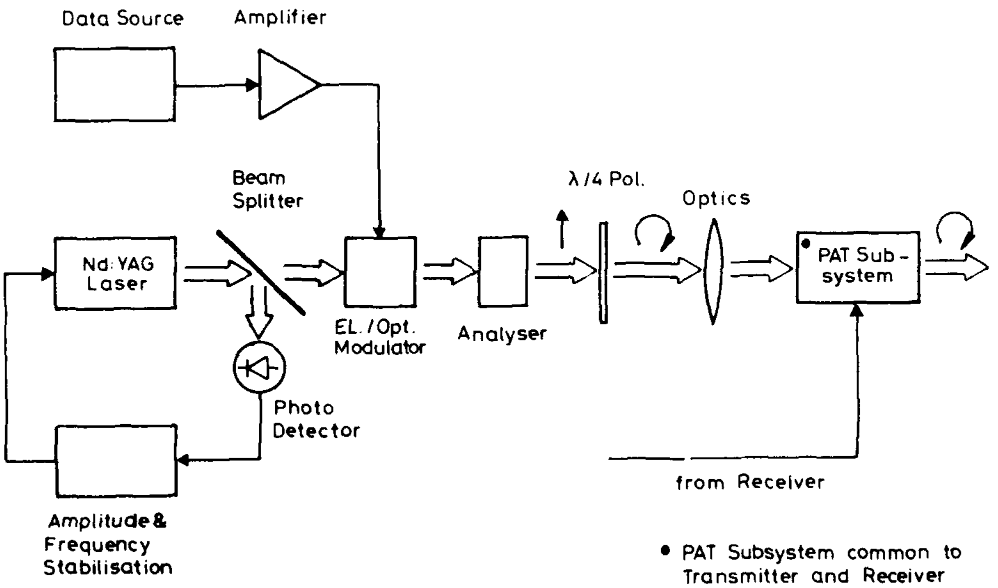


Fig.8.5.12 Nd:YAG LASER transmitter assembly (ASK)

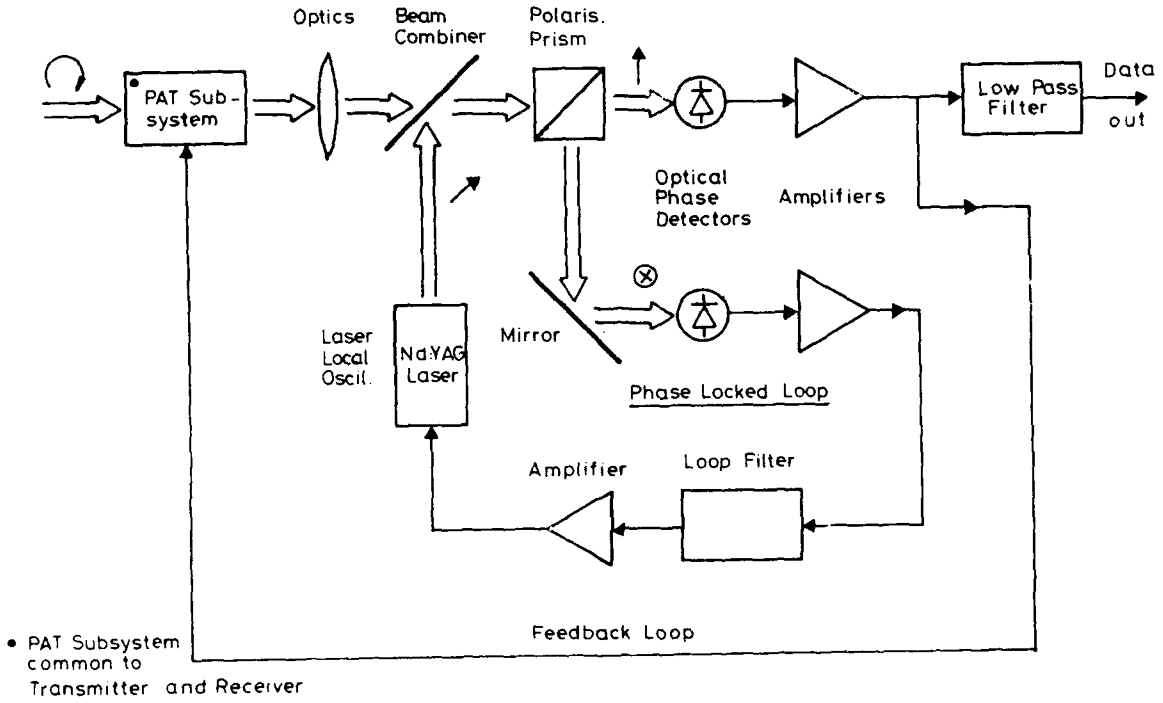


Fig.8.5.13 Nd:YAG LASER receiver assembly (ASK)

APPENDIX 8. 5A

COHERENT OPTICAL INTERSATELLITE CROSSLINK SYSTEMS

1. INTRODUCTION

High-capacity intersatellite communication crosslinks will allow more efficient and more reliable operation of future satellite systems. Crosslinks can achieve connectivity between terminals on opposite sides of the earth without expensive intermediate ground relay stations. Operating between a communication relay satellite and an observation satellite (weather satellite, earth resource satellite, or the space shuttle) intersatellite crosslinks also will elimanete the need for expensive world-wide ground tracking networks. Fig. 1 illustrates examples of possible intersatellite links.

Both microwave, e.g., 60 GHz, and optical communication technologies are candidates for future satellite systems. Microwave links with capacities larger than approximately 10 Mbps require antenna apertures and transmitter powers so large that it becomes difficult to integrate such systems on most host spacecraft. Future intersatellite links for satellite data acquisition networks or long-distance communication trucking systems may require capacities as great as several Gbps. Optical crosslinks offer the potential of operation at these high data rates because optical frequencies allow the use of very narrow transmit beams which can produce high recived signal levels with comparatively small antenna (telescope) packages. Furthermore, optical systems can provide very wideband communications channels and possess a high degree of immunity to interference, both intentional or otherwise.

A satellite communication package size and weight tends to be driven by aperture size, particularly for larger apertures. Fig. 2 compares antenna size requirements for different RF and optical technologies as a function of data rate. One of the earliest examples of a satellite crosslink was a 38 GHz link between the geosynchronous Lincoln Exyerimental Satellites 8 and 9 (LES 8/9) which were launched in 1976. 60 GHz crosslinks are now being considered for future systems. Recently, M.I.T. Lincoln Laboratory has been involved in the design and development of a space optical heterodyne system for the laser Intersatellite Transmission Experiment (LITE).

Two candidate optical technologies are direct (or incoherent) detection and heterodpne (or coherent) detection. Direct detection technology has been pursued for a longer time and thus is more mature. However, at the 0.8μm wavelength, heterodyne receivers can be 15 dB more sensitive than direct detection, thereby allowing for use of smaller apertures and/or lower transmit laser powe.. The reason for the sensitivity advantage is that heterodyne systems can provide near-quantum-limited noise performance whereas direction detection receivers are usually limited by photodetector excess noise, background noise, or amplifier thermal noise. GaAlAs lasers operating at 0.8 μm are compact, relatively rugged, and exhibit good prime-to-optical power conversion efficiency (roughly 10). Recent advances in GaAlAs lasers, particularly in the areas of power output and spectral purity, make them an important technology and system option for heterodyne crosslink systems.

Hetedrodyne optical communication receivers operate much like a conventional radio receiver. At the receiver, the incoming signal is mixed with a local oscillator (LO) laser and the beat signal at IF can be processed in much the same way as an RF signal. The heterodyne channel can be modeled as a classical additive white Gaussian noise channel, allowing theoretical-maximum receiver sensitivity can be realized with sufficient selectivity to permit operation even when the sun is directly in the receiver field of view (FOV). The elements of a heterodyne laser communication system are highlighted in Fig 3.

No experimental or operational laser communication systems have yet been flown in space. However, optical technology has evolved to the point where it is now feasible to consider such systems for space applications. The remainder of this Appendix will focus on the various system and technology issues related to implementation of heterodyne laser communication systems. The discussion will summarize the current state of the art in terms of key component technologies and system capabilities. Projections of future system and technology developments will also be made. Finally, the design and capabilities of possible flight packages will be presented.

2. KEY SYSTEM AND TECHNOLOGY ISSUES

The designer of a heterodyne laser intersatellite link is confronted with many system and technology issues. A complete discussion of all possible issues is beyond the scope of this paper. However, in this section, attention is directed to some of the most important areas.

2.1 Telescope\Aperture

As indicated earlier, aperture size is one of the key drivers of package size and weight. Fig. 4 illustrates the sharp growth in telescope and gimballed mirror weight with aperture diameter. This data is based on a scaling of the LITE design which has a 20 cm aperture with a 4 μrad diffraction-limited beam at 0.8 μm and uses a fixed telescope with a gimballed coarse-pointing mirror. As a rough rule of thumb, weight tends to scale as (aperture diameter)$^{2.4}$, although there is some deviation from this trend for smaller apertures.

In addition to increased weight, larger apertures produce smaller beamwidths and hence greater difficulty in pointing and tracking. Furthermore, the overall optical, mechanical, and thermal design becomes more complicated. Thus there is strong incentive for containing aperture size. Development of high-power lasers will permit realization of high-rate (> 1 Gbps) systems with modest apertures, e.g., 20 cm, in the future.

2.2 Laser Transmitter Technology

Perhaps the most important laser transmitter parameter is laser power. Single-mode lasers with 30 mW output are commercially available today. Active development is ongoing in the area of higher power lasers. As much as 1 W may be available in the future from coherent laser arrays and/or broad-area-devices. Wavelength-multiplexing of individual lasers is also a system option.

General requirements on lasers for heterodyne systems include: single-mode operation (spatial and frequency), sufficient tunability to track Doppler shifts which can be +/- 10 GHz in some applications, good frequency stability, narrow linewidth (low phase noise), long life (as much as 10 years for satellite missions), and stability of the laser electrical characteristics over time. Commercially produced GaAlAs lasers meeting these general requirements are available today, although further work is needed to fully characterize some aspects of the lasers long term aging characteristics. Linewidth must not exceed a small fraction of the data rate, depending upon the modulation scheme used, so as not to excessively degrade receiver

(*) Written by John E. Kaufmann and Vincent W.S Chan of M.I.T Lincoln Laboratory, USA.

demodulation performance [1]. Linewidth requirements for phase-shift keying (PSK) modulations are more stringent than for frequency-shift keying (FSK) or amplitude-shift keying (ASK). Present-day 30 mW lasers have been observed to have linewidths of 5 to 10 MHz which are compatible with data rates on the order of 100 Mbps or more if FSK is used.

With GaAlAs lasers, frequency modulation is realized conveniently by direct modulation of the laser injection current. Modulation speed is limited by the lasers intrinsic modulation bandwidth, which is typically a few GHz or more. Equalization of the lasers FM transfer function is required in FSK signaling. This is because unequalized laser diodes exibit a large enhancement in low-frequency drift of transmitted tones, depending upon the spectral content on the data sequence, and a resultant degradation in demodulatin performance. These effects can be mitigated significantly with passive equalization of the modulatin drive current [2]. Such techniques are simple to implement, do not require active electronic or optical components, and do not limit the inherent bandwith of the laser.

Phase modulation of laser diodes is also possible but requires use of external modulators or complex injection-locking techniques [3].

Heterodyne Reciver Technology

A strong receiver local oscillator (LO) laser, when, mixed with the incoming signal, provides gain to the signal. When the LO is sufficiently strong, the shot noise generated by the LO also dominates the noise of the subsequent receiver electronics and quantum-limited noise performance is attained. For heterodyne receivers in the 0.8μm region, this translates to a requirement of a few mW of LO power with photodetectors having a quantum efficiency of better than 50%. LO intensity fluctuations intrinsic to semiconductor lasers, can create excess noise in the receiver front end. However, an LO intensity-noise-cancellation technique has been developed using the dual--detector receiver shown in Fig.5 [4]. The excess noise currents from the two detectors are in phase with one antoher while the signal components are 180° out of phase. Thus subtracting the two detector outputs cancels the excess noise while completely recovering the signal. This arrangement also has the advantage of concerving LO power and enhancing system robustness in satellite applications in that the effects of increased LO intensity noise with laser aging can be minimized. In practice, a receiver noise floor within 0.5 dB of the quantum limit over a bandwidth of about 1 GHz has been achieved with this technique. [5]

Designing for both low noise and for wide bandwidth tend to be conflicting requirements. Demonstrations of heterodyne data receivers with bandwidths of several GHz have been reported, but in each case, noise performance has been considerably short of the quantum shot-noise limit. Thus design goals for future high-rate systems should stress bandwidth extension in conjuction with a near-quantum-limit noise floor.

The amplitude and phase distributions of the LO field on the photodetector must accurately match those of the received signal to yield high-efficiency heterodyne mixing. Spatial mode-matching of GaAlAs lasers with efficiencies within approximately one dB of the theoretical limit has been demonstrated recently [6]

Frequency locking of the receiver LO laser to the incoming signal using injection-current and/or temperature feedback to the LO will tract out laser center-frequency fluctuations and Doppler shifts between moving satellite platforms [7]. Fig. 6 depicts the configuration of a frequency tracting loop using current feedback, while and example of the stablization of the IF signal center-frequency at high SNR is illustrated in Fig. 7. Depending upon the satellite orbits, Doppler shifts at 0.8 μm. can be in excess of 10 GHz with rates greater than 10 MHz/s, although proper tracking system design can easily accomodate these shifts. Frequency locking is now routinely demonstrated with center-frequency stability of 2-4 MHz under realistic operating conditions, which permits efficient demodulation of FSK at rates of tens of Mbps or greater.

On the other hand, phase-locking of the LO to the incoming signal for PSK receivers is not yet practical because of the magnitude of phase noise in present-day lasers and the relatively low received signal power levels in intersatellite links [8]. Thus PSK signalling is not feasible unless the additional complexity of an external cavity laser is adopted. However, DPSK has considerably more relaxed phase stability requirements than PSK and will not require phase-locking at high data rates. Open-loop laser phase stability should be adequate for DPSK signalling at rates beyond 1 Gbps.

2.3 Spatial Acquisition and Tracking

The area of spatial acquisition and tracking presents some of the most significant design and implementation challenges for optical intersatellite systems. Because of the very narrow optical beamwidths, e.g., a few μrad, satellite terminals must acquire and track angle with great accuracy. In the acquisition phase, the receiver must search over an angular uncertainty area which is many beamwidths wide. This uncertainty arises from a number of sources, including ephemeris error, limitations in the accuracy of the optical pointing mechanisms, and probably most attitude control error. Attitude sensor errors, attitude control loop errors, and various mechanical alignment uncertainties limit attitude control accuracies to about one milliradian for geosynchronous satellites. Star trackers are capable of providing much greater accuracy, but should not be necessary to aid intersatellite link acquisition.

Each terminal must provide a beacon to aid acquisition by the other terminal. Search strategies include parallel search, zooming, or serial search [9]. Parallel search leads to the most rapid acquisition. Here, charge-coupled devices (CCD) or charge-injection devices (CID) are well suited to the role of acquisition detectors. Low-noise devices with as many as 1000 x 1000 pixels have been developed and can provide an acquisiiton FOV large enough (1 mrad x 1 mrad or more) to easily satisfy acquisition system requirements. Fig. 8 depicts a typical acquisition scenario, which begins with each terminal broadening its transmit beamwidth to illuminate the entire uncertainty region. Although the received power densities over a geosynchronous link distance will be low (on the order of 10^{-12} W/cm^2 for a 30 mW laser and a 1 mrad beamwidth), integration times in the CCD /CID can be chosen to build up a signal-to-noise ratio (SNR) large enough for signal detection. Acquisition times of no more than e few seconds are well within the capabilities of the present state-of-the-art.

The spatial tracker must track out local angular disturbances onboard the host spacecraft and relative translational motion of the other terminal, using the received beacon (which can also be the received data-modulated signal) as a reference. Tracking error generally must not exceed 0.1 beamwidths so as not to significantly degrade SNR and hence data demodulation performance at the receiver. Tracking system design requires a careful assessment of the dynamic environment of the spacecraft. Satellites can produce significant angular disturbances, varying widely in amplitude and frequency, from motion of solar arrays, RF antennas, attitude control system momentum wheels. Disturbances can be expected to be 100 μrad or more around 1 to 10 Hz, decreasing to a few μrad at 100 Hz and above. Closed-loop traching rejection and are achievable with current high-bandwidth mirror technology. Also, an open-loop point-ahead angle correction must be implemented and can be as much as 50 μrad, depending upon the link

Either direct-detection or heterodyne tracking systems can be employed. The heterodyne spatial tracker, like the heterodyne data receiver, has the potential advantage of greater sensitivity to

signal and relative immunity to background noise, including the sun.

2.4 Optical/Mechanical/Thermal Design

The optical subsystem design must satisfy a number of requirements: minimize throughput losses, provide high wavefront quality, maintain accurate beam pointing and alignment, provide transmit/receive isolation, and minimize optical feedback to the lasers. Performance must be maintained in an environment where the system is subject to a variety of thermal conditions and mechanical disturbance inputs from the spacecraft. Performance of the optical subsystem is clearly coupled to that of the mechanical and thermal subsysetms.

The mechanical design must first ensure physical survival of the package through launch. Furthermore, the design must provide a high degree of structural stiffness and stability so that critical optical alignments can be maintained and pointing accuracy is not degraded in the spacecraft operating environment. Where possible, isolation from onboard satellite disturbances should be provided. The mechanical design must also minimize weight.

The package can be subject to varying thermal scenarios, including full illumination by the sun, darkness, or transitions between the two cases, and varying conditions of generation of waste heat from electronics. The thermal design of the package must minimize and stabilize temparature gradients to minimize distortions or alignment changes in components such as mirrors and lenses.

2.5 Modulation/Demodualtion and Coding

Various forms of ASK, FSK, and PSK have been considered for heterodyne laser communication systems. The choice of modulation depends upon data rate, available bandwidth, laser linewidth, and modulator/demodulator implementation issues. Because of laser phase noise, only phase-incoherent schemes such as FSK are feasible at data rates below 1 Gbps unless an external cavity laser is used. Measured bit-error rate (BER) performance for a 100 Mbps FSK demonstration system is shown in Fig. 9. laser phase noise degrades FSK demonstration by two mechanisms: first, by reducing signal energy in the passband of the detection filter, and second, by introducing crosstalk in adjacent tone slots [1]. The onset of a BER floor in Fig. 9 is attributable to crosstalk. This floor can be significantly lowered by increasing the spacing between tones at the expense of greater overall signaling bandwidth, other techniques, including chip-combining at the receiver and use of error-correction coding can also be effective in mitigation linewidth effects. At data rates above 1 Gbps, DPSK becomes feasible, as discussed earlier.

Error-correction coding allows the channel to operate at higher error rates (typically 10^{-2} before coding) and yet provide a low BER (typically 10^{-6}) after decoding. Although some preliminary coding analyses have been performed [10], coding has not seen much application in demonstrations to date, primarily because of speed limitations in available decoders. Monolithic integrated decoding devices are highly desirable for space applications. However, the fastest present-day monolithic Viterbi decoder operates only up to about 15 Mbps. Further development in high-speed decoder technology is required before coding/decoding becomes practical at high rates.

3. FLIGHT PACKAGE DESCRIPTION

A block diagram of a laser heterodyne crosslink package is shown in Fig. 10. Four major subsystems are indicated: optical, electrical, transmitter, and receiver. The optical module contains the telescope, coarse pointing mirror, and other optical components associated with beam pointing and transmission. The electrical module provides conditioned electrical power and control processors. The receiver consists of a CCD acquisition system, spatial tracking detectors and electronics, the data receiver front end including LO laser, and demodulation electronics. Finally, the transmittter module contains the laser transmitter, its modulator, precision temperature and current controllers, and an autonomous diagnostics package which monitors laser power, wavelength, and modulation characteristics

Table I summarizes the gross characteristics of several systems using semiconductor lasers and spanning the data rate range of 10 Mbps to 5 Gbps for a 1 x synchronous-orbit link distance. The largest aperture size used in these examples is 20 cm. The 100 Mbps system is based on the LITE flight-experiment design done at Lincoln Laboratory. The systems for data rates above 1 Gbps are projections for the future based on development of higher power laser sources.

Table I
HETERODYNE LASER CROSSLINK SYSTEM CHARACTERISTICS

	10 Mbps	100 Mbps	1Gbps	5 Gbps
Source	30 mW	30mW	3x60 mW MPX/ 150 mW	800 mW
Aperture diam.	10 cm	20 cm	20 cm	20 cm
Modulation	FSK	FSK	3xFSK/DPSK	DPSK
Weight	190 lb	205 lb	280 lb/230 lb	250 lb
Power	210 W	225 W	335W/275 W	325 W

Fig. 11 gives estimates of package weight for three transmitter options: a single 30 mW laser transmitter, wavelength multiplexing, and a 1 W laser transmitter. The large weight growth above 100 Mbps for the 30 mW package results from the increase in aperture size. The 1 W system exhibits a significant weight advantage at high data rates because the higher power allows use of small apertures (≤ 20 cm). With wavelength multiplexing, extra weight for multiple receivers must be carried, although some overall weight savings relative to the 30 mW systems is realized. These estimates are based on extrapolations of existing system designs (primarily from the LITE program). Further technology development and application of hybrid circuitry and/or monolithic integration of the electronics may bring about additional weight reduction in the future.

Finally, an example of a link budget for a 100 Mbps heterodyne laser crosslink operating at 0.8 μm is shown in Table II.

Table II
100 Mbps HETERODYNE laser LINK BUDGET

Transmit laser (30 mW)	-15.2 dBW
Transmit optical loss	-4.0 dB
Transmit aperture gain (20 cm)	117.6
Space loss (40,000 km)	-295.6
Receiver aperture gain (20 cm)	117.6
Spatial tracking loss	-1.0
Receiver optical loss	-2.5
Photodector Q.E.=0.8	-1.0
Spatial mode match loss	-1.5
Detected power	-85.1 dBW
Detected photons/s	101.1 dB-Hz
Receiver front-end noise	-1.0 dB
Implementation loss	-2.0
Receiver sensitivity (photons/bit @ 10^{-5} BER)	- 12.0
Data rate (100 Mbps)	- 80.0
Margin	5.6 dB

4. SUMMARY AND CONCLUSION

Laser communication becomes an attractive alternative to microwaves for interasatellite links operating at data rates of 10-100 Mbps and higher. Heterodyne technology with

semiconductor lasers has evolved to the point where it is now feasible to implement heterodyne systems for ~ 100 Mbps with a single-laser transmitter of modest power (30 mW) and with small apertures (20 cm). Technology advances, particularly in the area of higher power laser sources, will extend system capabilities beyond 1 Gbps in the future.

5. REFERENCES

1. L.L. Jeromin and V.W.S. Chan, "M-ary FSK Performance for Coherent Optical Communications Systems Using Semiconductor lasers", IEEE Transcations on Communications, Vol. COM-34, No. 4, April 1986.

2. D.Welford and S.B. Alexander, "Equalization of Semiconductor Diode laser Frequency Modulation With A Passive Network", Electronics Letters, Vol. 21, No. 9, April 1985, pp. 361-362.

3. Kobayashi and T. Kimura, "Optical Phase Modulation in an Injection Locked AlGaAs Semiconductor laser", IEEE Journal of Quantum Electronics, Vol. QE-18, No. 10, October 1982, pp. 1662-1669.

4. G.L. Abbas, V.W.S. Chan, and T.K. Yee, "A Dual-Detector Optical Heterodyne Receiver for Local Oscillator Noise Suppression", IEEE Journal of Lightwave Technology, Vol. LT-3, No. 5, October 1985.

5. G L Abbas et al. "Near Quantum-Limited Operation of a GaAlAs laser Heterodyne Communication System", OSA Conference on lasers and Electro-optics, Baltimore, MD, May 1985.

6. K.A. Winick and P.Kumar, "Spatial Mode Matching Efficiencies for Heterodyned GaAlAs Semiconductor lasers", IEEE Journal of Lightwave Technology, Vol. 6, No. 4, April 1988.

7. F.G. Walther, S.D. Lowney, and J.E. Kaufmann, "Frequency Tracking for Heterodyne Optical Communications Using Semiconductor lasers", Paper FC3, OSA Conference on lasers and Electro-optics, Washington D.C., June 1983.

8. V. Chan, L. Jeromin, and J. Kaufmann, Heterodyne laserCOM Systems Using GaAs lasers for ISL Applications", IEEE International Conference on Communications, June 19-22, 1983, Boston, MA.

9. P.Van Hove and V.W.S. Chan, "Spatial Acquisition Algorithms and Systems for Optical ISL", IEEE International Conference on Communications, Paper E1.6, Boston, MA, June 19-22, 1983.

10. J.E. Kaufmann and L.L. Jeromin, "Optical Heterodyne Intersatellite Links Using Semiconductor lasers", IEEE Global Telecommunications Conference, Paper 28.4, Atlanta, GA, November 26-29, 1984

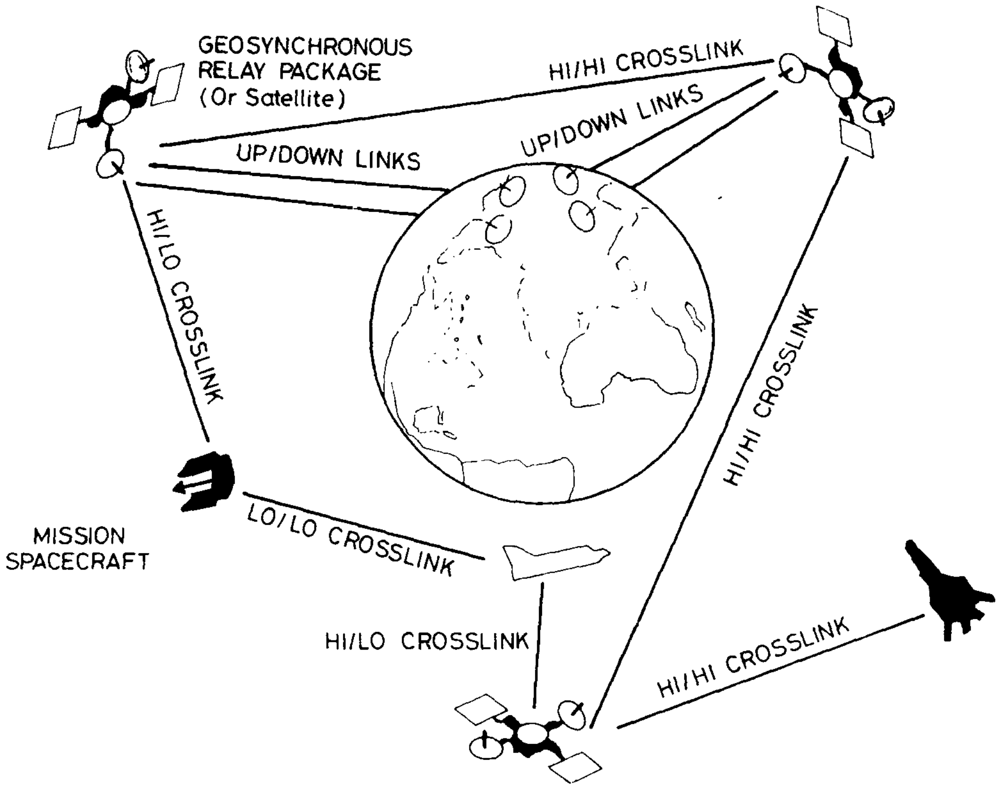


Fig.1 An internetted space architecture

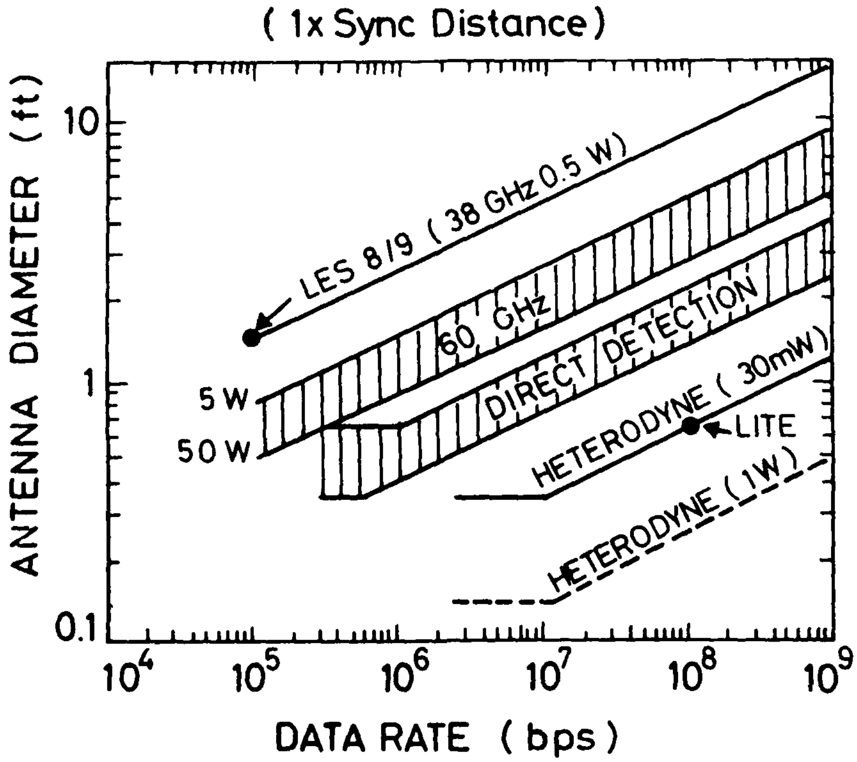


Fig.2 Comparison of crosslink systems

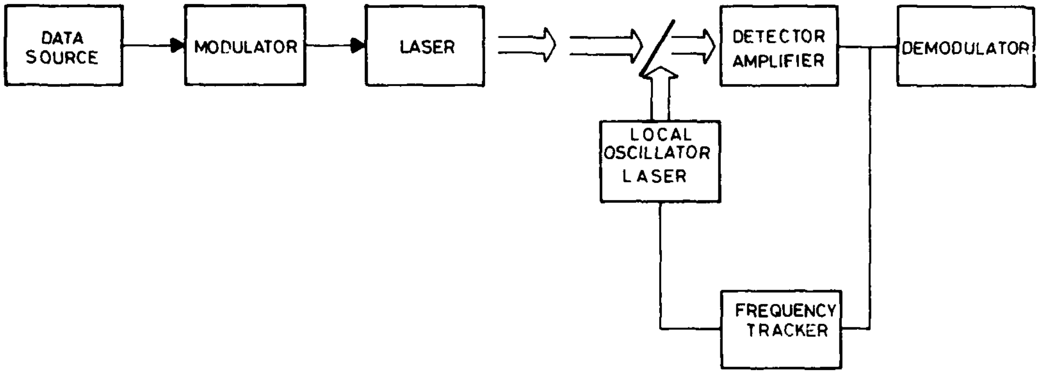


Fig.3 Optical heterodyne communication system block diagram

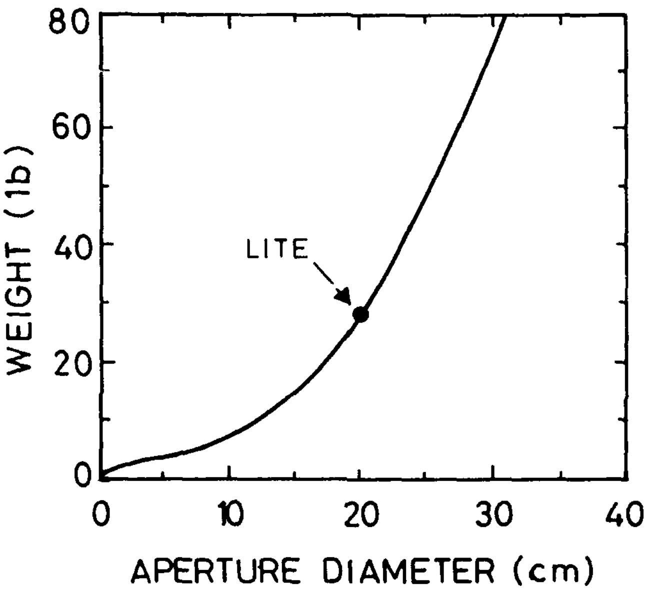


Fig.4 Telescope/gimbal weight

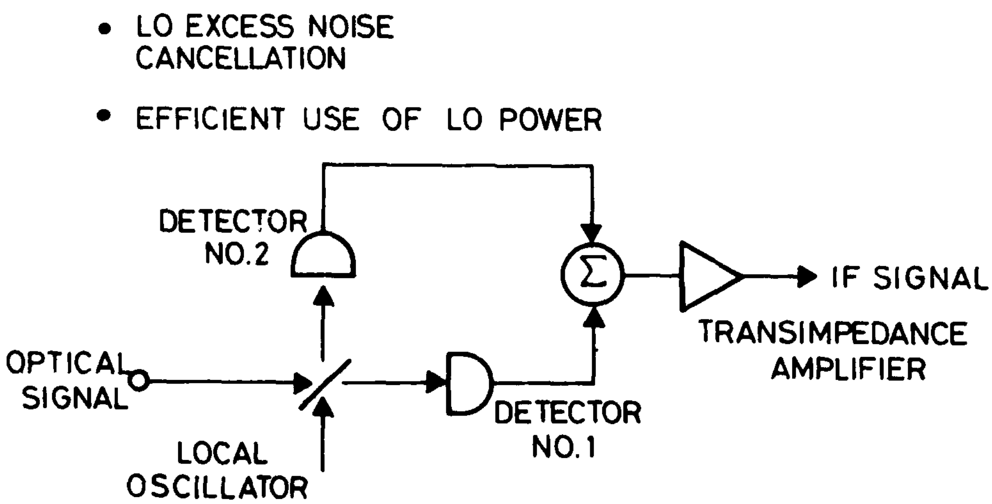


Fig.5 Receiver front end

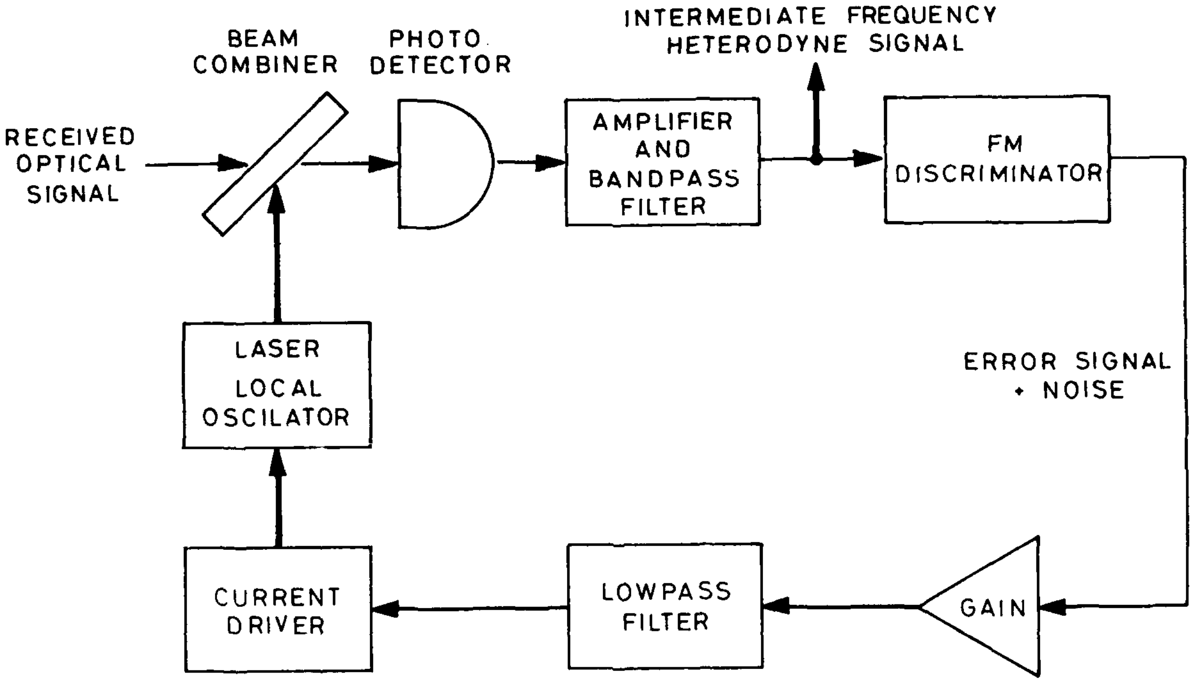


Fig.6 Receiver

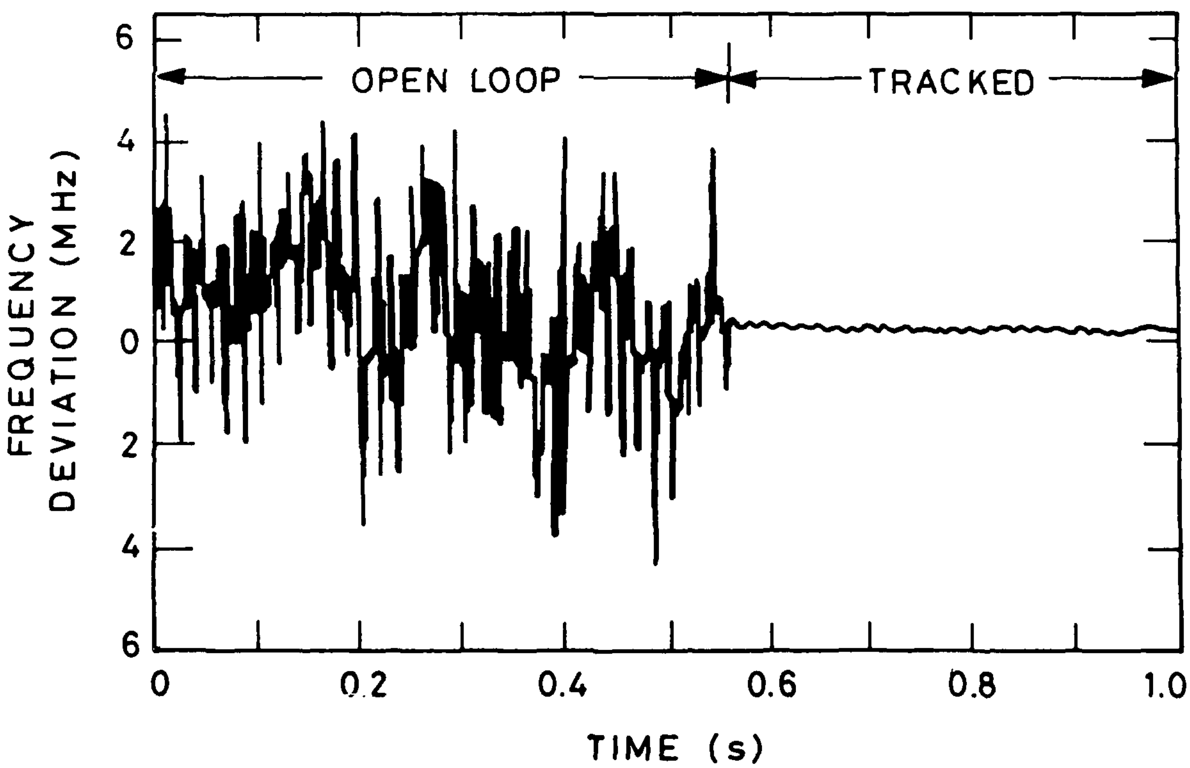


Fig.7 Frequency deviation

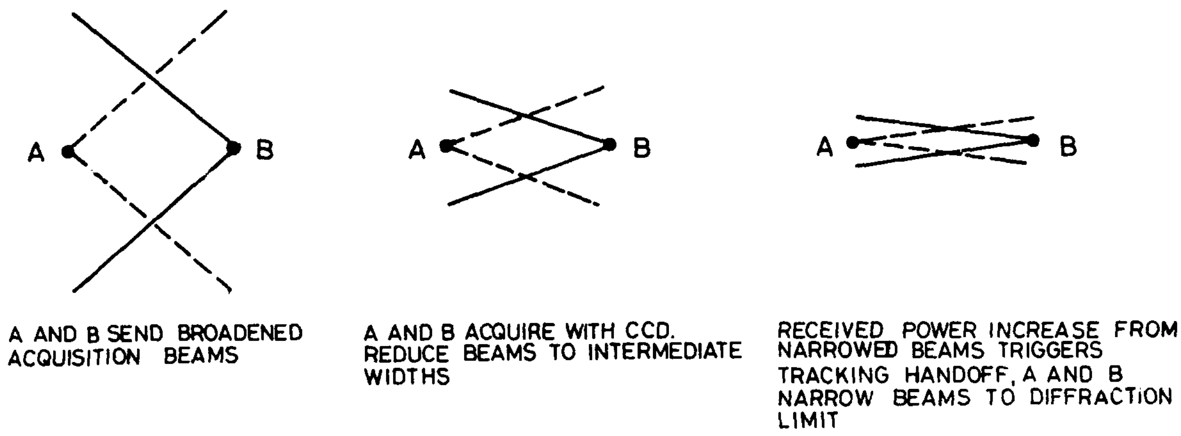


Fig.8 Acquisition sequence

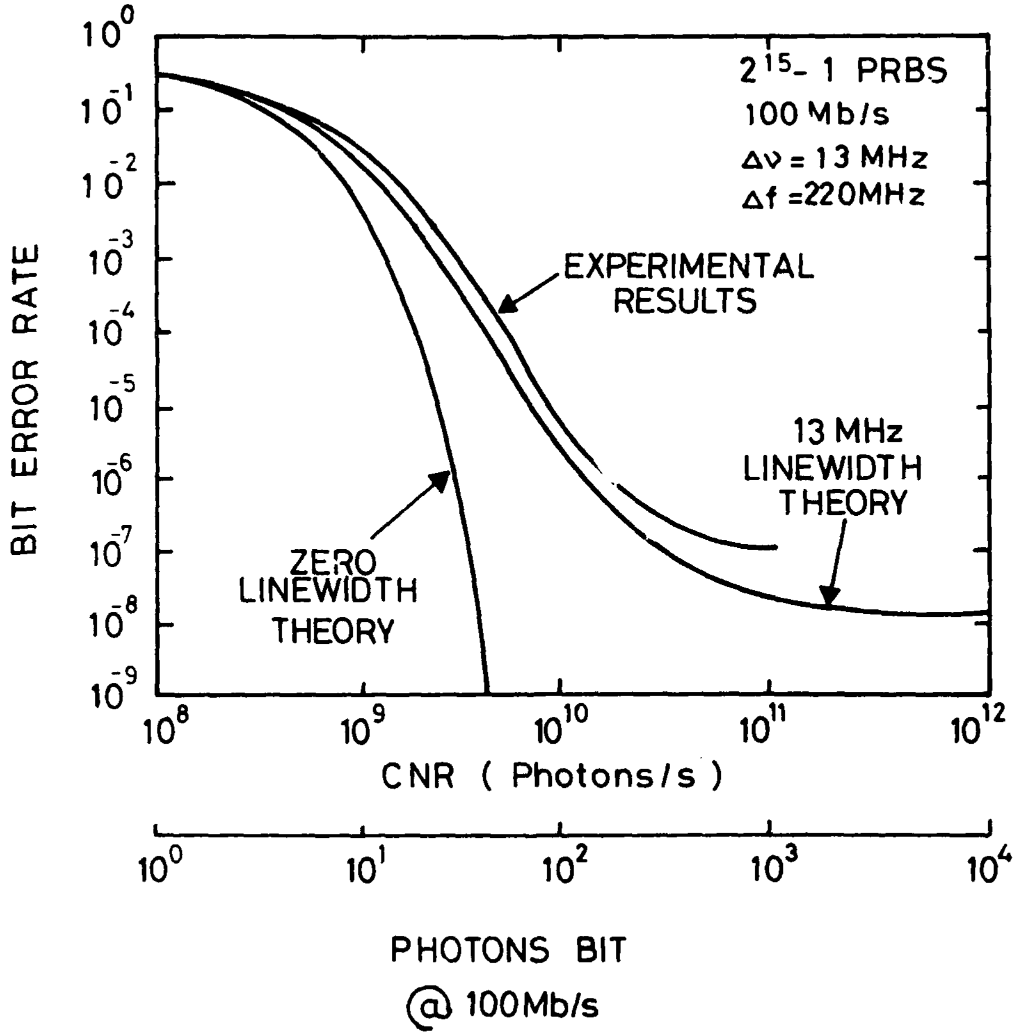


Fig.9 Performance of binary FSK

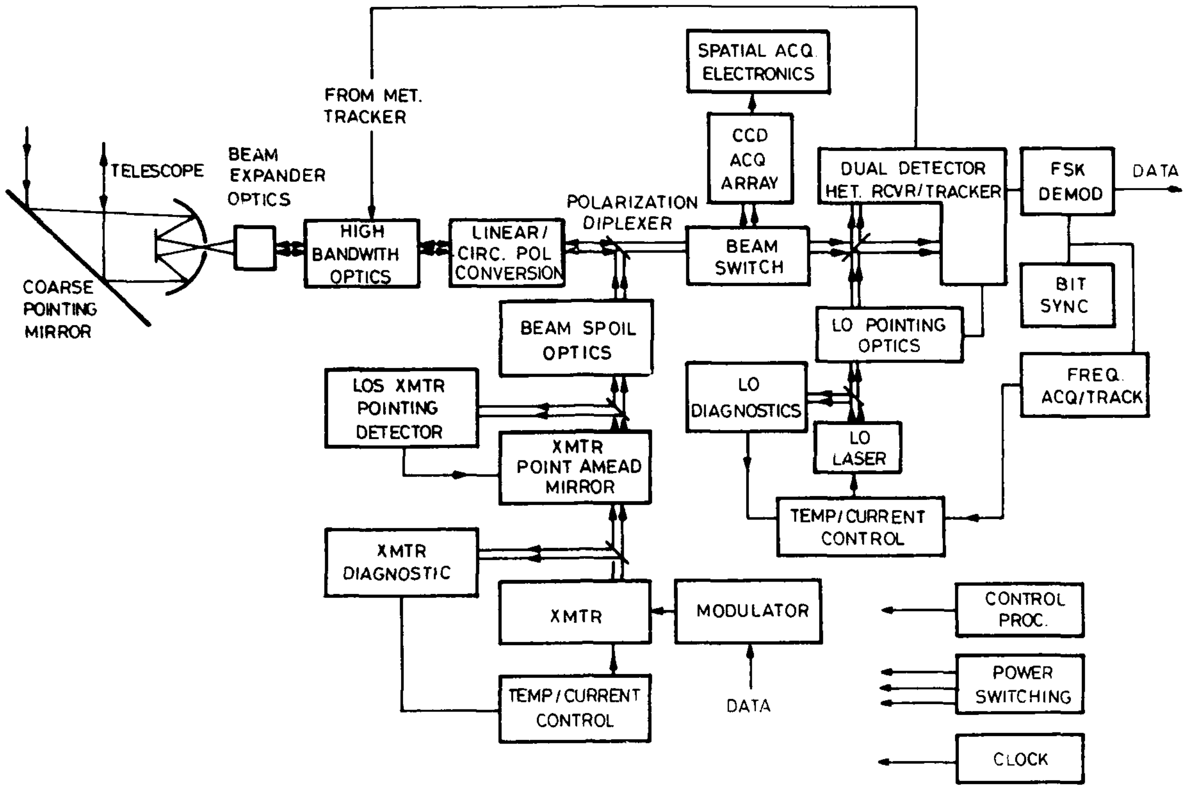


Fig.10 LASERCOM heterodyne crosslink package

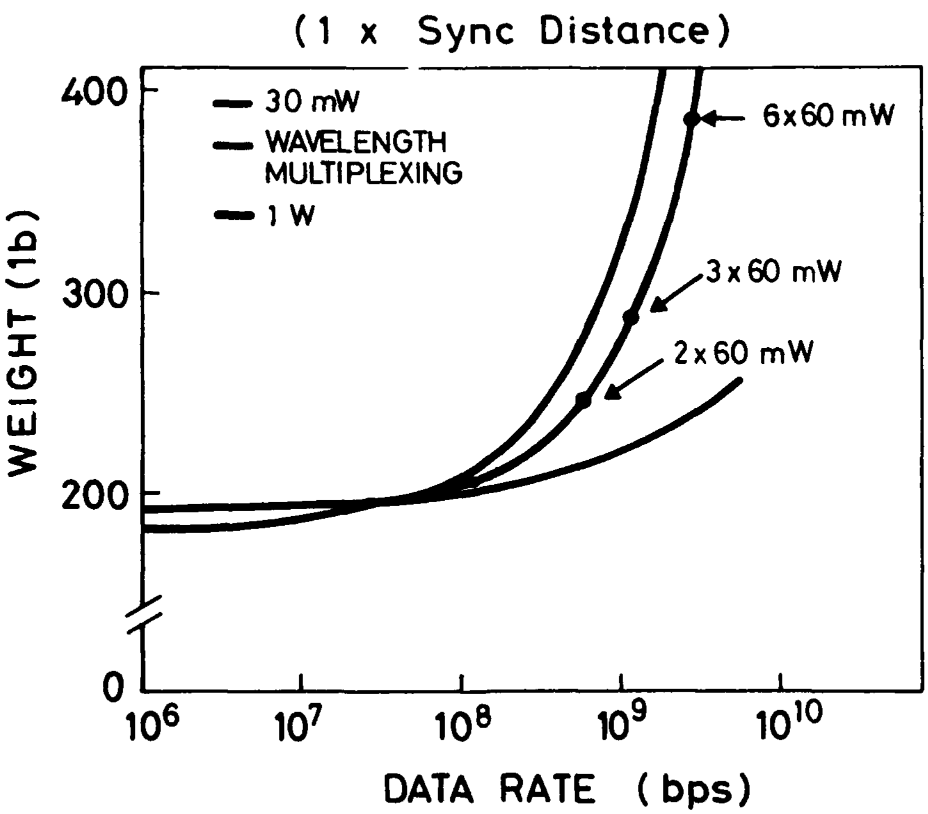


Fig.11 Heterodyne LASERCOM crosslink package weight

8.6 LASER COMMUNICATIONS TO SUBMERGED SUBMARINES

8.6.1 Requirements

NATO has a requirement for communications to submarines at depth in high northern latitudes. Communications to submerged submarines has historically been considered as a strategic requirement for controlling the arsenal of submarine launched nuclear missiles. However, in recent years surface fleets have become increasingly vulnerable to guided missiles (eg, the Faulklands). The role of tactical submarine warfare has assumed a much greater significance. Correspondingly, requirements for submarine communications now extend into the tactical arena. The efficiency of tactical submarines would be multiplied manyfold if they could receive target information at depth and while engaged in operations. This eliminates the need for them to break off operations periodically to surface in order to receive instructions, which reduces their effectiveness and makes them more vulnerable.

So, interest in submarine communications is high for both strategic and tactical reasons. This working group has, therefore, investigated the feasibility of including a submarine communications capability in future NATO satellite systems. The concept is to use laser signals.

8.6.2 The Concept Of Submarine Laser Communications

Communications to submarines is based on the fact that sea water has only two windows of transparency in the entire electromagnetic spectrum. One is at extremely low frequencies, ELF, and extends into the VLF range.

Sea water is a conductive medium with a certain characteristic skin depth. This skin depth permits electromagnetic waves to penetrate to depths of several tens of meters at frequencies around 50 Hz decreasing to several meters in the VLF part of the spectrum. This low frequency window is the basis for several existing submarine communications systems such as TACAMO and Sanguine.

These low frequency systems have some fundamental limitations. They can support only very low data rates. They require large antennas to be deployed and towed by the submarine, which increases its vulnerability. Also, depending on the exact frequency used, the submarine must ascend to within a certain distanece of the surface to receive its messages. All of these characterictics seriously limit the operational utility of ELF and VLF submarine communications systems.

The other window in the electromagnetic spectrum is the so-called "Jerlow minimum", the prominent andrelatively narrow dip in attenutation (Figure 8.6.1) that occurs around 6 x 10^{14} Hz, or a wavelength of 500 nanometers, which is in the center of the visible band. In this window, attenuation drops from hundreds of dB per meter to a few tenths of a dB per meter. The Jerlow window is the basis for the SLCSAT concept. The prospect of communicating to submarines through the Jerlov window using laser beams from satellites has been under consideration since the late nineteen sixties.

8.6.3 The Channel

The laser propagation path from satellite to submarine includes free space, clouds and sea water. The mode of propagation in the latter two media is by scattering of the photons from particles: moisture droplets in the case of clouds and suspended particulates in the case of sea water. Scattering is characterized by beam attenuation and angular spreading and by the smearing of the signal in time.

Figure 8.6.2 illustrates the scattering phenomena in a cloud. Thick clouds attenuate the signal by about 10 dB, which is not significant compared to the by many tens of dB attenuation it will suffer once it enters the water. The predominant cloud effect is speading in time. An infinitely narrow microseconds by multiple scattering effects as it traverses the cloud. Propagation by multiple scattering is essentially a process of photon diffusion through the medium.

The other cloud effect is to spread the beam in angle. This is not very significant in view of the relatively short distance between clouds and water. Its only result is to increase slightly the already-large footprint of the laser beam on the ocean's surface.

Propagation through the sea water also occurs by multiple scattering. The predominant effects here are attenuation and angular spreading. By the time a pulse of laser light reaches a submarine at several hundred meters depth, it has been attenuated by absorption in the sea water, dispersed in time over several microseconds and spread in angle over essentially the upper hemisphere. Total losses in the sea water can be many tens of dB.

8.6.4 The Receiver

The primary noise source in the SLCSAT channel is the sun, which peaks in intensity at the same frequency as the signal (Figure 8.6.3). Sunlight would completely swamp the signal if any conventional means of optical filtering were used. The feasibility of SLCSAT hinges entirely on the invention of a special type of optical filter that achieves exceedingly narrow bandwidths over very large acceptance angles.(In conventional optical filters narrow bandwidth and wide angle are mutually exclusive requirements.) This special filter, known as the atomic resonance filter (ARF), uses the cesium atoms in a vapor to capture the signal photons over a wide angle of incidence and convert them to longer wavelength red photons, which are picked up by conventional photo sensors. Only photons at exactly the right wavelength can undergo this conversion, hence the exceedingly narrow bandwidths (fractions of an angstrom). This filter was invented at the Lawrence Livermore Laboratory based on earlier work a at the Lincoln Laboratory.

8.6.5 Performance

With the ARF narrow band, wide field of view filter, it is possible to think of a SLCSAT system architecture that permits communications with submarines at several hundred meters depth with burst data rates of several hundred baud under worst case conditions (daylight, cloudy, poor sea water quality). The optimum modulation scheme is pulse position modulation, which requires a laser transmitter that is capable of a pulsed mode of operation.

The average data rates are considerably less than the burst rate. The laser beam footprint covers a spot several kilometers wide on the ocean surface. The data rate within the spot is at the burst rate, however, the spot is scanned in angle to "paint" a large segment of ocean surface with the same message. This is so as not to reveal the submarine's location.

The average data rate is, therefore, the number of bits in the message divided by the time it takes to paint the entire coverage area.

With the satellite in synchronous orbit, the laser requirements are for several tens of watts of optical output power with pulse repetition rates of several tens of Hertz.

8.6.6 The Laser Transmitter

If the ARF receiver is a crucial component of the SLCSAT concept, then an equally critical component is the laser transmitter. It must be at exactly the right frequency with an average power output and pulse energy that are appropriate and which can be launched in a satellite and operate reliably in orbit for several years unattended. Unfortunatelly, there are very few lasers which ideally meet all these conditions (in fact, the number is approximately zero). The most promising candidate at this time is a xenon chloride (XeCl) gas excimer laser whose ultraviolet output is downshifted to match the cesium resonance frequency by means of a Raman scattering cell. The output power, pulse rate and wavelength of this laser are ideally suited to the SLCSAT architecture. For this reason it has been carefully researched. Versions of this laser system have been tested in aircraft-to-submarine links with spectacular results. Effective command and control of submarine forces has been uemonstrated in large scale naval maneuvers and even communications through the polar ice cap have been demonstrated with airborne lasers of this type.

Unfortunately, however, the XeCl laser system is not very suitable for use in satellites in its present state of development. It is inherently a large and complex device (Figure 8.6.4) and contains such items as mechanical pumps and circulators, corrosive gases, high energy electrical discharge currents and lead vapor operating at nearly 1000° C. There are a number of associated reliability and lifetime issues. The high peak currents in the electrical discharge cause serious electromagnetic compatibility problems with nearby equipment. The necessary resources for develooing a flight qualified version of this laser have not been made available.

A search for a solid state alternative laser is underway at present. Frequency doubled neodymium YAG lasers operate in the green part of the spectrum and would penetrate the sea water well enough.

However, there is no receiver equivalent to the cesium ARF that match the doubled YAG frequency. An earlier search several years ago yielded no other potential solid state laser condidates and so far none have been identified by the newer searches. This is not to say that a breakthrough is impossible, but the situation would seem to require just that in order for progress to resume.

8.6.7 Conclusions

This working group has concluded, after carefully reviewing the available information, that SLCSAT must be viewed as a possible future option pending some as yet unforesen technical breakthrough.

The breakthrough might be of a nature that would lead to resuming the development and flight qualification of the XeCl laser system. If this were the case, the size of the laser transmitter would almost certainly dictate that the SLCSAT part of the NATO SATCOM architecture be a dedicated satellite or at least a dedicated platform of a cluster architecture.

On the other extreme, the breakthrough might come in the area of solid state lasers. If this were the case one would hope for a SLCSAT laser transmitter capable of being integrated into a satellite with other communications equipment. This would certainly be the desired outcome from NATO's perspective.

At this time, however, this working group can not predict whether a breakthrough will occur or what its nature would be if it did occur. SLCSAT is, therefore, relegated to the status of "maybe" in the hierarchy of future NATO capabilities.

8.6.8 Bibliography

[8.6.1] T.F. Wiener et al., "The Role of Blue/Green Laser Systems in Strategic Submarine Communications" IEEE Trans. on Com., Vol. COM-28, No.9, Sept.1980.

[8.6.2] J.B. Schultz, "Navy, DARPA Evaluate Blue-Green Lasers as Communications Link With Submarines", Defence Electronics, Nov. 1983.

[8.6.3] F.C. Painter, "Submarine Laser Communications Best of TACAMO, ELF?, Jan. 1986, DS&E.

[8.6.4] R.S. Clark, "Getting the Laser Word to Subs", Photonics Spectra, March 1988.

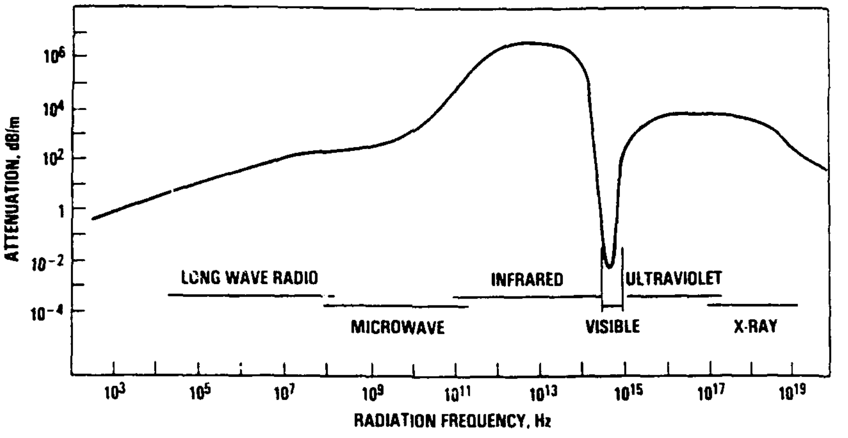


Fig.8.6.1 Attenuation of sea water in decibels per meter

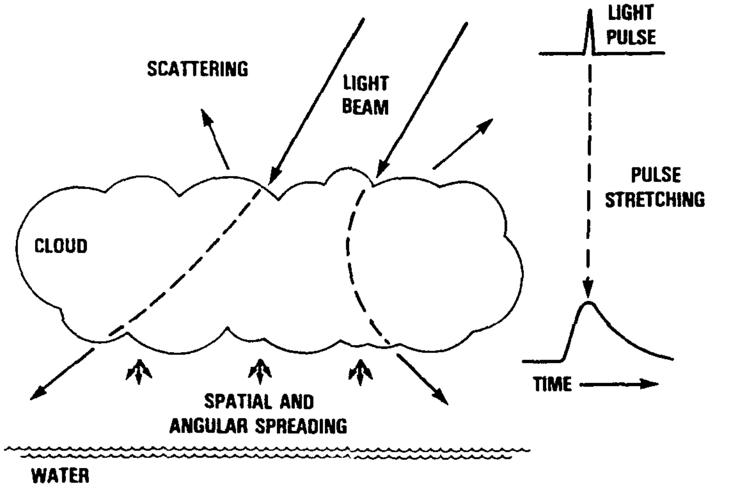


Fig.8.6.2 Scattering phenomena in a cloud

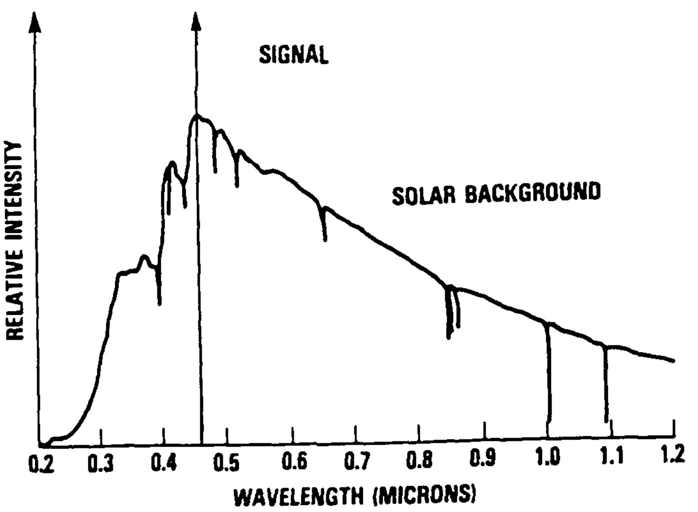
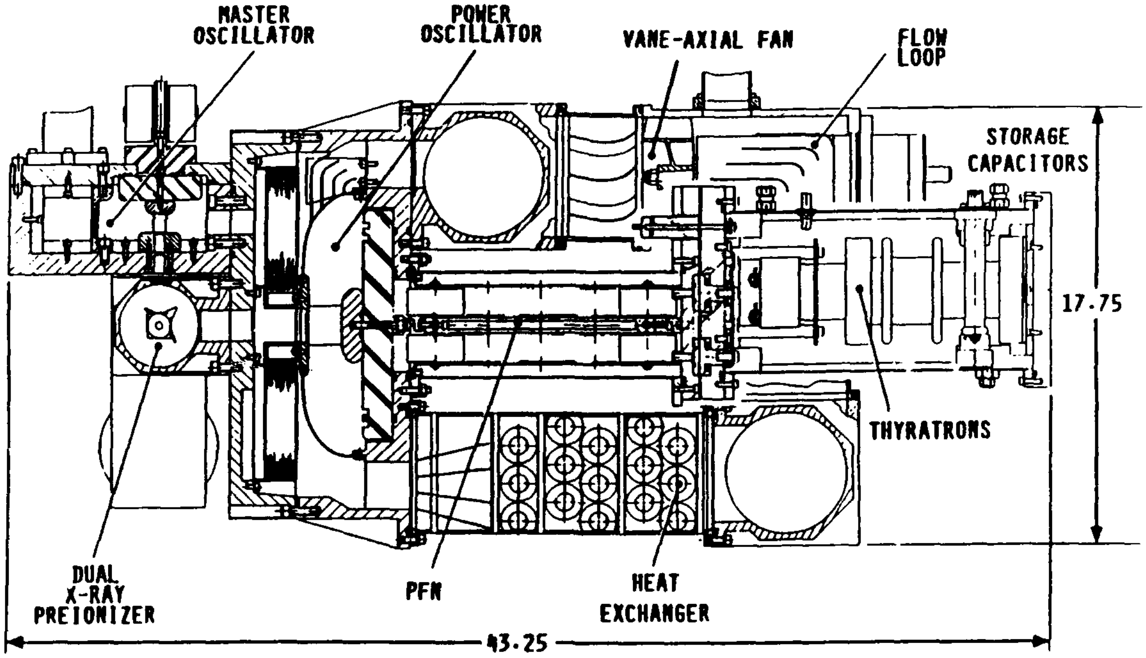
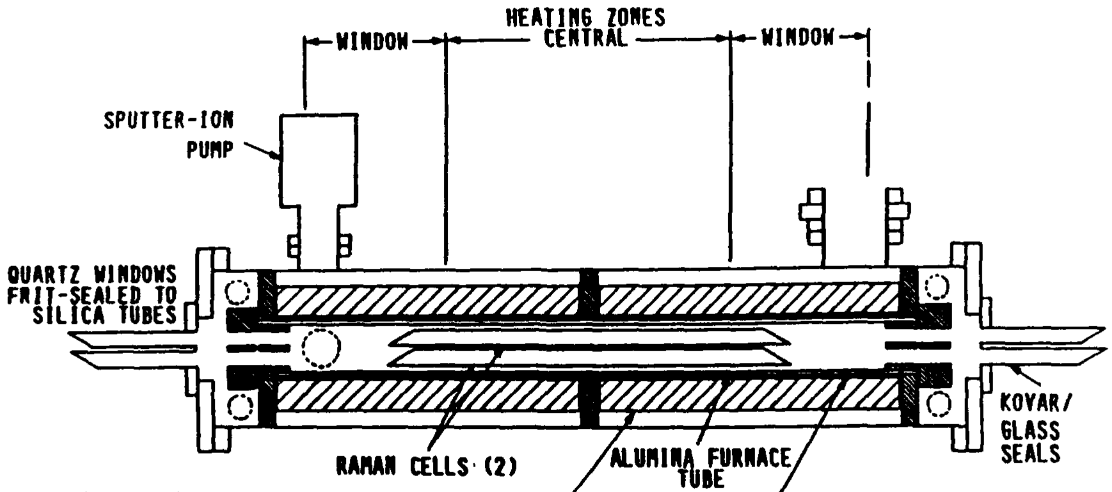


Fig.8.6.3 The solar spectrum is the primary noise source in a SICSAT channel



LASER CROSS-SECTION



RAMAN CELLS

- ALL SAPPHIRE
- 60 cm LENGTH
- Pb^{208} FILL, NEON BUFFER
- 1850°C FRIT

RAMAN CELLS

Fig.8.6.4 The Xenon Chloride (XeCl) LASER transmitter

8.7 ADVANCED MATERIALS FOR SPACECRAFT

8.7.1 Introduction

There have been a number of breakthroughs in materials and materials processing techniques which have led to the development of highly sophicticated spacecraft sub-systems and electronic equipment. The "conventional" materials, such as aluminium alloys have the advantage of being well characterised from the point of view of design data and their performance under differing conditions. They have been used for many years in both aircraft and spacecraft. The experience gained in that time has resulted in a high degree of confidence in them and the structures from them. Advances with these conventional alloys has come from processing innovations such as superplastic forming and diffusion bonding.

However, advanced materials are finding more and more applications in new designs. This is particularly the case with polymer composites containing carbon or aramid fibres, clean materials (with low outgassing) and several new types of metal alloys [8.7.1 & 8.7.2].

Advanced projects such as the Space Station and Hotol will continue to require advanced materials with high strength (particularly at high temperatures), stiffness, wear resistance, vibration damping, thermal control, resistance to atomic oxygen, etc.

The prospects for new materials are encouraging and new processes enabling material charcteristics to be modified are also becoming available from industry.

8.7.2 The Environment and Material Considerations

The environment to which spacecraft may be subjected, can be classified as follows:

- Natural space environment (LEO/GEO)
- Deliberate physical threat.

8.7.2.1 Natural Environment

Natural space environmental effects are increasingly better understood and their activity is dependent on the operational orbit (LEO or GEO). The effects are:

(i) U.V. radiation
(ii) Charged particles
(iii) I.R. radiation
(IV) atomic oxygen
(v) vacuum
(vi) temperature cycles

In general terms, the gradual improvements seen over recent years in the performance of materials in these environments will undoubtedly continue and essentially no quantum technology leap is required to provide extended life-times.

The advent of re-useable launcher systems and their ability to retrieve exposed components for examination has done much to extend our understanding of the LEO environment and its effect on materials. Typical of these are CTE changes in polymers exposed to charged particles and, of course, the extremely aggressive nature of atomic oxygen, an effect which prior to Shuttle flights and subsequently to LDEF had been grossly underestimated.

The super-imposition of atomic oxygen resistance may require some modification to the coating/blanket/second surface mirror technology of materials typically developed to control the radiative factors above. Thus coatings based on the Si-O combiniation seem to be particularly resistantwhether in the glassy or rubbery form. Fluorinated polymers, especially those based on aromatic (benzene ring) spinal molecular chains also showed early promise but existing data is sometimes contradictory.

The structure and protective coatings for communications equipment, e.g., antenna dishes, phased arrays will also require onward development of atomic oxygen resistance when these are required to operate in LEO. Currently, the problem is solved by making communication reflectors etc from metal or metallising non-conductive versions. However, frequency selective, patch array antennas requiring non-conductive substrates, will require protection of the dielectric surface without having any significant effect on the R.F. performance

Materials displaying improved outgassing and temperature cycle (microcrack) resistance seem to be developing by exploiting those materials which we currently know as thermoplastics although modified thermosets are mirroring those developments. Some evidence that polymers mixing the structure of both thermosets and thermoplastics may provide an answer.

No hardwared has yet been retrieved after long term exposure in GEO. Current assumptions are based on data received from instrumented GEO spacecraft and on the lergely successful lifetime performance of communications satellites. A retrievable experiment is perhaps long overdue

8.7.2.2 Deliberate Physical Threat

At present, it is diffucult to quantify accurately at what levels the following threats might be imposed on spacecraft.

(i) Laser
(ii) Kinetic energy
(iii) Particle beams
(iv) Nuclear burst effects

As discussed in Section 5.5 it is understood that laser threats are limited to LEO, i.e., sufficient power generation only from earthly station and energy dispersion of beam such that geostationery vehicles will receive only relatively low energy densities. In LEO, the damage mechanism to the spacecraft would be thermal overload, thereby causing damage or even vaporisation of surface materials.

The protection from such laser attack must depend on an ability to:

(i) deflect the hazard
(ii) absorb the input energy

The first of these implies the use of reflective filters and/or metallised films to provide a reflective capacity preferably with greatest efficiency in the frequency range in which the laser operates.

The second factor requires materials with a high thermal capacity (thermal mass), conductivity and an ability not to degrade at elevated temperatures. It is essential to understand that despite the ability to reflect a proportion of the input energy a large amount will need to be absorbed. This implies the possible use of following listed metal matrix composites (MMC), alloys and composites.

(C, SiC in Al, Ti(MMC))

(Al-Si-Fe-Zr-Ti alloys/Ni-based alloys, rapidly solidified))
(C-C, SiC-SiC ceramics composites)
(Polyimides + others polymer composites)

The major problem here is trying to reduce mass by using low density products but mass is part of practical measure of thermal capacity, therefore all solutions will be heavy.

Kinetic energy devices use particles of greater mass than micrometeoroids and at typical impact velocities the release of energy is tremendous.

The problem here is fairly obviously to reduce the kinetic energy of the particle. In essence, there are two factors which may be used to reduce the energy level, i.e., to reduce the mass by smashing the particle into smaller pieces/vaporisation or to reduce the velocity factor. The former technique was exploited on GIOTTO where the first shield, a metal sheet, reduced the mass by converting sufficient energy into heat to vaporise/pulverise the particle into smaller pieces to be absorbed, some distance behind the initial shield, on a thick Kevlar composite panel.

Deliberate threats are likely to involve larger masses and, therefore, a primary shield is likely to be ineffective and all energy must be absorbed for the single particle. This, therefore, demands that such a shield, even made from the most energy-absorptive material must have considerable bulk and consequent mass.

A particle beam weapon would suffer similar restrictions to a lasser weapon in terms of it size, and energy requirements ard, therefore may be less effective for high altitude satellites. The beam would have to consist of neatural particles such as neutrons because the earth's magnetic field will deflect charged particles excessively.

A recent American study has concluded (Ref. Sec. 5.5) that the technical problems accociated with aparticle beam weapon system are even more severe than those associated with laser weapons, therefore this type of weapon will be unlikely to be a dominant threat in the frame of interest but if a requirement arises, one can deploy shields made from neutron absorbing materials like C, B, Be.

The primary damage producing phenomena due to nuclear burst are: x rays, Gamma rays, neutrons, radioactive debris, thermal pulse and EMP. X rays in sufficient quantities produce surface damage. Gamma rays and neutrons penetrate the spacecraft and can cause damage to solid state devices. All of these nuclear effects are relatively well known and understood. There are design techniques and material processes known as 'hardening' which reduce these effects and can be applied to a greater or lesser degree depending on the level of hardening required. Implementing these techniques will have a significant effect on satellite mass.

8.7.3 New Material Concepts

The property goals for basic research of aerospace materials are frequently listed, and while their order may reflect the importance attached to them in a particular context, it is unlikely that the list would contain any surprises. What would be surprising is if the "low cost" parameter was not at the bottom of each list and yet it is probably the cost element more than others which dictates the eventual pace of material development. Low cost is perhaps also the most difficult of the listed parameters to achieve. Recent study indicates that improvements in other properties such as stiffness are won only against an exponential cost penalty.

Earliest material development was based on the combination of various elements and the subsequent management of the resulting system, i.e., alloying and heat treatment.

This type of development, still pertinent today, has resulted in a range of materials:

i) Metals
ii) FRP composites
iii) Ceramics
iv) Polymers

which have themselves become the building blocks of a new family of products: enginereed composites.

8.7.3.1 Metal-Based Systems

Advantages:

i) isotopic properties
ii) good thermal conductivity
iii) low outgassing
iv) good resistance to radiation/atomic O
v) good electrical conductivity
vi) high strength/ductility

Disadvantages:

i) generally high density, except Al, Mg, Be
ii) limited but not poor maximum service temperatures
iii) low-stiffness (specific)

So far, Metal Matrix Composites have found no structural applications on ESA spacecraft but have done so on other vehicles such as the Space Shuttle. Metal matrix composites are clearly an important future development area. Aluminium and magnesium reinforced with carbon fibres are obvious candidates materials for such a development effort. Metal matrix composites have a number of desirable properties including zero moisture absorption, stable in the presence of atomic oxygen and electron, proton and UV radiation and high through-thickness thermal conductivity. The main disadvantages of these types of materials and their high cost and the difficulty of processing them, both of which should be prime candidates for research in order to relieve such problems.

Another difficulty with carbon fibre/aluminium composites for high precision structural applications has been thermal strain hysteresis and residual dimensional changes resulting from matrix yielding during thermal cycling.

Most composite development has so far relied on bringing the already formed components together for subsequent combination. Such a process has to overcome imcompatibilities of structure, mechanical properties, thermal expansion and even excessive chemical affinity of one component for the other.

For future development, another approach might attempt to grow the reinforcement within the matrix with the hope of achieving greater compatibility. Such techniques have already been exploited in self-reinforcing polymers but extension to metal-ceramic combinations may require the development of high temperature/high pressure techniques coupled with reinforcement growth control via say, magnetic or electrical fields.

8.7.3.2 Fibre Reinforced Plastic Composites

Spacecraft have continued to make use of thermosetting and, so far to a much lesser extent, thermoplastic resin-based composites. Such materials as uni-directional and woven carbon fibre, aramid and glass fibre composites are procured in both sheet ond roll form. Cutting is carried out by hand or by machine and, in the case of the latter, the cutting requirements can come direct from the geometry of parts developed on a CAD system. Guidelines on such composite applications are now well established, although important aspects such as design allowables, the significance of defects, damage tolerance and

failure criteria need further evaluation. Up to now this has not been a major problem, but will become important for Ariane-5 and Hermes manned missions.

Some carbon fibre/thermoplastic matrix composite have very good mechanical property profiles and glass transition temperatures above 200°C and are, therefore, favoured for high temperature applications. These so-called 'aromatic polymer composites' include fibre reinforced PEEK (Polyetheretherketone) developed by ICI. However, it should be remembered that the 'Achilles heel' of PEEK is the hysteresis effects of thermal cycling and the limitation this puts on its use

8.7.3.3 Ceramic-Based Systems

Leading contender for onward development is porbably SiC.-SiC composite.

- has excellent thermal properties;
- process improvements-reducing defects and thus statictics of catastrophic failure;
- superb oxidation resistance;
- oxidation resistant-including atomic O;
- electrically non-conductive

8.7.3.4 Polymer-Based Systems

For space applications polymer development might embrace areas such as:

i) improved maximum service temperatures for both reinforced and unreinforced roles e.g., polyimides and beyond.
ii) reduced tendency for microcracking.
iii) improved mechanical performance e.g., self-reinforcing polymers.
iv) extend range of materials with radiation resistance. atomic oxygen resistance etc
v) improve conductivity (electrical/thermal).
vi) creep resistance thermoplastic.

8.7.4 Discussion and Conclusions

Much work remains to be done, particularly to counter the effects of atomic oxygen in low earth orbit, and already several materials and protective coatings have been evaluated. Oxidation of material was not a significant problem with the relatively short Space-Shuttle missions, but it could have an unacceptable effect on CRFP. Kapton and silver should these common materials be selected for exposed elements of the Space Station. Several new processing techniques have already been discussed, but possibly one deserving further evaluation is the method developed for the unique tailoring of alloys and protective finishes using atom-by-atom deposition from metal vapours. "functionally gradient materials".

The effects of radiation on materials, particularly non-metallic composites also need further study. However, there are problems with radiation testing in that it is not easy to reproduce the conditions found in orbit and it is very expensive.

Work on investigating the effects of atomic oxygen on materials is under way and so far indicates the need to protect some non-metallics with metal coatings where long service life is required.

Manufacturing techniques also require attention as do design methods and design data acquisition, particularly relating to composite materials.

8.7.5 References

[8.7.1] Dauphin J. "Selection and Control of Materials for Space Applications" ESA Journal, Vol 8, pp.53.

[8.7.2] Dunn B.D. "Materials and Processes Division" ESA Journal 1987, Vol.11 ESTEC, Noordwijk, The Netherlands

BILIOGRAPHY

[8.7.3] "Spacecraft Materials in a Space Environment", Toulouse 8-11 June 1982 ESA SP-178

[8.7.4] "Third European Symposium on Spacecraft Materials in a Space Environment", ESTEC Noordwijk The Netherlands 1-4 October 1985

[8.7.5] "Advanced Composites Technology", Papers presented at The Conference on Advanced Composite Technology El Segundo California 14-16 March 1978

[8.7.6] Lenoe E.M., Oplinger D.W & Burke J.J. (Eds)," Fibrous Composites in Structural Design". Plenum Press 1980

[8.7.7.] "Practical Considerations of Design, Fabrication and Tests for Composite Materials", AGARD Lecture Series No 124 AGARD-LS-124.
11-12 October 1982, Oporto, Portugal
14-15 October 1982, London, UK
18-19 October 1982, Ankara, Turkey

[8.7.8] "ICCM VI & ECCM 2", Second European Conference on Composite Materials 20-24 July 1987. Imperial College of Science and Technology, London, UK

[8.7.9] Hansen J.S., and Tennyson R.C. "Simulated Space Environmental Effects on Fibre-reinforced Polymeric Composites", AIAA-83-0589, AIAA 21st Aerocpace Sciences meeting, 10-13 January 1983, Reno, Nevada

[8.7.10]Gerharz J' and Schulz D "Literature Research on the Mechanical Properties of Fibre -Composite Materials-Analysis of the State of the Art", Volume 1. Royal Aircraft Establishment. Library Translation 2045. August 1980.

[8.7.11]King R.L. "Statistical Derivation of Design Data for Composite Materials Computed Aided Design in Composite Material Technology", Proceedings of the International Conference. University of Southampton, UK.

[8.7.12]"Materials in Aerospace", 2-24 April 1986 Proceedings, Volumes I and II Royal Aeronautical Society

[8.7.13]Hoggatt J.T., Hill S. G. and Johnson J.C. (Eds.)," Materials for Space--- The Gathering Momentum", 18th International SAMPE Technical Conference, Red Lion Inn, Seattle, Washington 7-9 October 1986

[8.7.14]Williamson, J.R., "Advanced Materials for Space Structures", 41 st Congress of the International Asronautical Federation, Oct. 6-12, 1990, Dresden.

8.8 SUPERCONDUCTIVITY

8.8.1 Materials

The subject of superconductivity as it impacts electronics in general is treated in Appendix 8.8A.

During the active period of the working group there has been a remarkable progress in superconductors with high critical temperature.

Commercial supreconductors in use today operate at 4.2° Kelvin, the temperature of liquid helium at atmospheric pressure. The materials and the technology required are well understood. The main obstacle to take these superconductors into use in communication satellites is the bulky equipment required to refrigerate to liquid helium temperatures. The metal niobium (Nb) can be used in form of superconducting films and wires. At microwave frequencies the RF surface resistance determines the performance of the superconductor. The surface resistance increases with frequency for a niobium film at 4.2° Kelvin. But even at 50 GHz it is at least one order of magnitude lower than copper conductors cooled to 4.2° K.

The new materials with high critical temperatures can be operated at liquid nitrogen temperature, 77° K, which will mean a large reduction in the complexity of the cooling equipment. Materials with T_C as high as 125° K has been demonstrated. Beyond 2000 even higher T_C may be practical and the cooling required may be achieved simply by radiation shielding and exposure of the equipment towards the free space temperature. Such possibilities would have a major impact on several on-board subsystems, the impacts perhaps most fundamental in making more AJ signal processing possible. Some possibilities are to be mentioned in the next section. The high T_C superconductors are a family of materials where super-conductivity depends on copper oxides. There is a growing understanding of the superconducting mechanism in these materials and it seems clear that the conducting mechanism is highly anisotropic. To use superconductors in analogue or digital signal processing equipment, superconductive films are preferable. Deposition of thin films with acceptable microstructure for superconductivity is a research topic where substantial progress is reported [8.8.1]. Currently primitive microstrip resonators and transmission lines have been demonstrated.

8.8.2 On-Board Applications

a) Power subsystems.

The power subsystem operates at DC and low frequencies where superconductivity is nearly perfect. On the other hand, also the current technology with conventional metal conductors gives power subsystem with high efficiency. The power subsystem of a satellite is also distributed over a large area which makes refrigeration expensive or may be impossible. Application to the power subsystem may therefore depend on a breakthrough in high T_C materials.

b) Antenna

This subject is treated generally in Appendix 8.8B. To provide a sufficient radiation aperture to generate the EIRP, the receive sensitivity and the angular resolution required reflector antennas may employed. The current density of reflector antennas is generally low and ohmic losses represents, even with conventional materials, only a minor degradation. For reflector antennas superconductivity seems to be of minor interest with the possible exception of details in the feed horns. Super-conductivity may be of more interest to phased array solutions where each separate antenna element has a low gain and where the ohmic losses in the radiating element and in the distribution network may be a major cause of degradation. The cooling arrangement may be a problem, however, if this can be solved through efficient refrigeration or very high T_C temperatures, superconductive phased arrays may be an attractive solution in particular for the nulling receive antenna.

c) Mixers.

Superconducting mixers are established technology for radio astronomy at frequency above approximately 50 GHz [8.8.2]. The noise temperature obtained is close to the theoretical limit at the lower end of the frequency range. For a satellite uplinks where thermal noise is received from the surface of the earth, it seems that high electron mobility transistors will give the performance required and be much simpler.

d) Resonator filters.

For satellite communications in general and in particular for AJ communication the input filter bank (see. Fig. 8.2.1) is an important system component. Multiresonator filters can be designed from well established theory. The major degradation is caused by the ohmic losses in the resonating cavities. Superconductivity can at a given resonator quality factor support much higher current density which can be exploited to miniaturize the radio frequency filter bank. With conventional conductors cavity resonator with higher order low loss modes are utilized. With superconductivity these can possibly be replaced by stripline resonators and at the same time reduce losses. The input filter bank is a system component with limited mass/volume which makes cooling more feasible. If thin ilm high T_C superconductors become available an uplink filter bank becomes much more possible for future NATO satellites. Nb superconductors at liquid helium temperature can also be a possibilitiy dependent on progress in refrigerating equipment capable of liquid helium temperatures.

e) Chirp Fourier Transformers

The Chirp Fourier Transformer described in Sec. 8.2 can also be implemented by using electromagnetic chirp lines. With conventional conductors the transmission line losses prevents useful components to be made. However, with superconductors useful combinations of device bandwidth and dispersive delay can be realized without going to excessive bulky devices. From the system considerations given in Sec. 8.2 it is reasonable to discuss a system where the two GHz uplink band is demultiplexed into 20 slots. Each of these slots can then be processed further by the SAW CFT discussed in Sec. 8.2. To obtain 20 slots, a chirp line time bandwidth product (TBP) of 200 is required. This TBP has been demonstrated with Nb superconductors. [8.8.3] The required selectivity (100 MHz per slot) can be reached by using say. 35 nsec dispersive delay. Also this value has been realized with Nb superconductors. High T_C components would of course simplify refrigeration of the system. Super-conducting Chirp Fourier Transformers are believed to be highly interesting system components for future NATO AJ satellites. The working group recommends NATO to follow the progress in this area closely. The key to this application is high T_C superconducting films and/or more efficient refrigeration equipment.

f) Digital technology.

Superconductivity can enhance the speed and compactness of digital hardware. Several on-board subsystems can in principle benefit from such a development. However, it seems that the digital filter banks in the AJ processor is by far most important due to the high processing load in this equipment. If major break

throughs in digital speeds are attained, DSP may be a candidate to replace the SAW chirp fourier transformer in the system approach indicated in Sec. 8.2.

8.8.3 References

[8.8.1] M.R. Beasley: "High-Temperature Superconductive Films" Proceedings of the IEEE, Vol. 77, No. 8, Aug. 1989 (pp. 1155-1163).

[8.8.2] P.L. Richards and Q. Hu: "Superconducting Components for Infrared and Millimeter-Wave Receivers", Proceedings of the IEEE, Vol. 77, No.8, Aug. 1989 (pp. 1233-1246).

[8.8.3] R.S. Withers and R.W. Ralston: "Superconductive Analog Signal Processing Devices", Proceedings of the IEEE, Vol. 77, No. 8, Aug. 1989.

APPENDIX 8. 8A

THE IMPACT OF SUPERCONDUCTIVITY ON ELECTRONICS(*)

The first announcements of high T_C superconductivity last year were very quickly followed by press promises of many wonderful applications. Among small scale ones, they ranged from superfast computers to new brain scanners, from eatrhquake predictors to microwave detectors. Moreover the expectation was that these would become realities long before the high field magnets and levitated vehicles which might characterise the corresponding 'large scale' revolu on.

In this paper, some of these propi.ecies will be assessed. To do this it will be necessary to review foundation of these predictions, which is that liquid helium superconducting electronics is in fact already a mature and successful technology. For each aspect of the field, we can then estimate how it can be extended to higher temperatures. Materials development, particularly of thin film structures, will control the rate of progress and we will discuss some of the difficulties. Finally we will show that the press forecasts of 'early delivery' are already proving correct, by mentioning some rudimentary magnetic sensors which are already operational at temperatures up to 90°K.

Applications (both possible and actual) of superconductivity can conveniently be classified under headings which describe the property exploited.

1) Zero resistance
2) Meissner effect-magnetic screening
3) Giaever tunnelling
4) Josephson RF devices
5) Josephson logic
6) Superconductor-semiconductor hybrids
7) Josephson magnetic sensors-SQUIDs
8) Macroscopic quantum chorence

1) Zero resistance

Already used in striplines and transmission lines, and in high Q microwave cavities. Extension of former to high T_C likely and has already been demonstrated for pulse transmission. More general interconnects (need high current density) are also likely. Cavities will be more difficult because of inter-grain losses.

There are likely new applications to zero resistance antennas. Low losses in high power short dipole transmission, conversely noise advantages in reception.

2) Meissenr Effects

Perpect diamagnetism in low fields (up to 0.01-0.05T), because of persistent surface screening currents. Already much used at 4.2K for magnetic shields and 'field trapping' in tubes. Screening likely to be important at the new temperatures, both at the tiny scale, for low intuctance 'ground-planing and on chip screening and on the large scale, in structures such as field free rooms for medical monitoring with SQUIDs. Trapping gives an 'adjustable' permanent magnet.

3) Giaever Tunnelling

Uses S-I-S or S-I-N diodes where S,N are 0.5μm films of superconductor or normal metal. I is 2nm of insulator. Electron tunnelling given current-voltage behaviour which is very non-linear on scale of electron energy gap Δ (typically a few millivolts in 'old' superconductors, corresponding to far infrared). Already appli to ionising particle detectors, Raman spectrometers, three terminal 'transistor' and (most prominently) to low noise microwave detectors and mixers (with gain at 100GHz). Mixers much used in radio astronomy, now up to 400 GHz (750GHz). Likely extension to 1THz.

New superconductors might go to 10THz (larger gap). Use in astronomy, radar applications, etc possible. But tunnel junctions will be very difficult to make (see later).

4) Josephson RF devices

The Josephson effect is observed in 'weak link' structures between two superconductors, which are point contacts, microbridges, or S-I-S thin film tunnel structures. Several Josephson effects exist (see also 5 & 7): here we consider the AC effect whereby junctions emit, and are particularly sensitive to, radiation of frequency ν where $\nu = 2eV/h$, V being the bias voltage. This relation is exact, corresponding to 484MHz per microvolt of bias; it is now the basis of the Josephson volt whereby worldwide standards are established in terms of standard frequencies rather than chemical cells. Chips with series arrays of >5,000 junctions are now available, with an direct output voltage of 1-5V and an accuracy of 1 in 10^9. By eliminating the need for liquid helium. High T_C superconductivity will make such standards cheaper and more accessible. The larger Δ value will mean fewer series junctions per volt of output, simplifying manufacture.

The Josephson RF effect is also used in mixers and detectors, and in attenuation standards.

5) Josephson logic

Pairs of Josephson junctions in parallel form a cell, which when biased with a current of order 10μA, can be switched very rapidly (~10ps), from V=0 to V=2Δ, using a magnetic control line. A ultrafast logic family is then available which has zero standing power, and dissipates 10^{-17}J per interrogation. This is the basis of the Josephson computer, abandoned by IBM, but still vigorously pursued in Japan. It is also used in a number of less complex structures such as fast A-D converters (8bit 1GHz, with 16 bits suggested), and picosecond sampling circuits.

The new superconductors are not promising for large scale Josephson logic, though they may be useful for cache memory and some of the smaller systems. They may switch as fast, but their dissipation will be larger and overall they will probably be inferior to FETs.

6) Superconductor-semiconductor hybrids

Offer the greatest scope for truly novel devices, including new fast computer systems. Only marginal progress has been made at helium temperatures, since semiconductor carriers are almost always frozen out, though superconductivity has been induced and controlled in n-channel FET structures at 4.2K using the proximity effect (establishment of superconductivity in neighbouring normall metals). But at 77K, where semiconductors are still 'active', the opportunity to combine them with the gap-related and phase-coherent properties of superconductors in hybrid structures creates possibilities for novel 3-terminal devices and even optical devices.

7) Josephson magnetic devices

SQUIDs (Superconducting Quantum Interference Devices) combine the sensitivity of Josephson junctions to magnetic field with the most subtle feature of superconductors, which is the

(*) Written by Gordon B. Donaldson of Department of Physics, University of Strathclyde, Glasgow, Scotland.

extensive phase coherence of their electron waves. A loop of superconductor containing one or two Josephson junctions shows a current-voltage response which is periodic in the magnetic flux applied to the loop. The period is the flux quantum ($\emptyset_o$ = h/2e = $2x10^{-15}T\text{-}m^{-2}$) For a $1mm^2$ loop this implies a period of $2x10^{-9}T$; appropriate feedback electronics makes it possible to subdivide this period, producing magnetometers of extraordinary sensitivity. In a commercial SQUID the noise limit on the subdivision is about $2x10^{-4}\emptyset_o\text{-}Hz^{-1/2}$; the current research limit is 100 times better than this.

SQUID magnetometers are usually operated together with superconducting input coils configured as field gradiometers, and in such modes can detect tiny fields from local sources even in the presence of the earth's background field. The best sensitivities quoted are order 1 fT(10^{-11} of the earth's field, about 5 orders of magnitude superior to a flux gate magnetometer) and extend down to DC. Applications of such sensitivity have included geophysics, archaeology, submarine detection, non-destructive testing, NMR detection, ELF communication, and esoteric topics like searches for gravitational waves and magnetic monopoles. But the most important application for SQUIDs seems likely to be to biomagnetism, and particularly to magnetoencephalography-real time measurement of the magnetic field outside the brain due to electrical activity inside. Such measurements (the fields are of order 10-100fT) can be used to analyse and localize brain features such as sites of epileptic activity. There is now a strong push towards developing arrays of up to 100 thin film SQUIDs for multichannel biomedical systems.

Could SQUIDs based on high T_C materials and run at 77K or higher fit into this activity? Inevitably the noise will be worse for thermal reasons. However a typical sensitivity for ideal tunnel junction SQUIDs looks likely to be about $4 \times 10^{-5}\emptyset_o\text{-}Hz^{-1/2}$, better than the current limit for commercial (point contact) devices. In fact rudimentary SQUIDs based on simple thin film and bulk structures have already achieved $1 \times 10^{-3}\emptyset_o\text{-}Hz^{-1/2}$, and though not very reliable they hold out great promise for a wide range of new applications of SQUIDs. Many of these would be 'mobile' -eg NDT, mineral surveying, etc, for which the disadvantages of liquid helium, though trivial in the laboratory, become overwhelming in the field.

8) Macroscopic Phase Coherence

The electronic phase coherence in a superconductor extends over the entire (macroscopic) dimensions of the material. In some senses a ring of superconductor is a quantum object, like a single hydrogen atom but scaled up 10^{14} times in all directions. Certain consequences, like discrete energy levels, have already been demonstrated in the laboratory. The reader is invited to speculate on the possibilities if room temperature superconductors become reality.

High T_C Materials

Most electronic applications of superconductors rely on thin films; realising high T_C applications will usually need high quality S-I-S tunnel junctions as well. Considerable progress is being made in the science of thin films of YBaCuO and by sputtering, e-beam, or MBE on to special substrates, oriented thin films with good current carrying capacity ($\sim 10^6$ $A\text{-}m^{-2}$) are being made. Unfortunately water damages the films, and techniques for lithography are rather limited, but progress is being made.

Tunnel junctions are more challenging. The coherence length in these materials is only about 1nm and is a measure of the interfacial sharpness which will be necessary at boundaries between the ceramic superconductor and the oxide insulator. This will call for very sophisticated fabrication procedures.

Conclusion

Superconductive electronics was already a successful, expanding technology before the new materials appeared. Many of the existing applications will benefit from the availability of high temperature superconductors, and there are prospects for several completely new devices as well.

SQUIDS and their applications

Introduction

SQUID types

With SQUIDs (Superconducting Quantum Interference Devices). there there have been available for several years low frequency (DC to 500MHz) electromagnetic sensors of extreme sensitivity. In special cases the energy resolution has reached the limit set by Heisenberg's uncertainty principleapproximately 10^{-34} Joules in approximately a 1 sec bandwidth. Many applications, several of interest to MOD, have been generated. The attractions of operation 77K, especially for mobile applications, make it necessary to consider (i) the likely performance degradation due to higher thermal noise (ii) progress in the fabrication of High T_C SQUIDs and (iii) the implications for existing and possible new applications.

SQUIDs are flux to voltage converters which rely on the periodic magnetic properties, on a scale of the flux quantum, $\Phi_o = 2 \times 10^{-15}T\text{-}m^{-2}$, of Josephson weak-links. They are of two types. The RF-biased version is the easier to make, needing only one Josephson junction which can be a point contact; it is the basis of commercially available SQUIDs, which are bulk structures. The DC SQUID, on the other hand, requires two matched junctions (usually tunnel structures): although most of the problems have now been solved, it has proved the more difficult to develop. because its fabrication requires micron scale thin film technology. This type is the only one suitable for large scale integration.

Noise-4.2K and 77K (theoretical)

The performance of most SQUIDs intended for instrumental applications can be characterised by the flux noise Φ_N exhibited when they are used in conjunction with the appropriate room temperature electronics. As Table 1 shows, at 4.2K good DC SQUIDS are about 50x better than RF SQUIDs in the white noise regime (above about 0.5Hz). DC SQUIDs continue to be superior at lower frequencies, but for both types the noise degrades according to a 1/f relation. This is an important consideration for applications such as MAD, navigation and geophysics, and research continues into the origin and suppression of 1/f noise.

Theoretical estimates (see Table 1) of likely noise performance at higher temperatures can be made for each type of SQUID on the assumption that the noise processes which dominate at 4.2K will also determine the behaviour of High T_C devices. It seems that operation at 77K is likely to degrade RF SQUID noise performance by only a factor 4, while if the quality of High T_C tunnel junctions can ever be raised to that of present classical ones 77K DC SQUIDs will still be several times better than current 4.2K commercial (RF) devices.

TABLE 1

Type	White Noise ($\theta\text{-}hz^{-1/2}$) 4.2K (typical)	77K (predicted best)	77K (achieved)
RF (commercial)	1×10^{-4}	4×10^{-4}	7×10^{-4}
DC	2×10^{-6}	4×10^{-5}	10×10^{-4}

The results are encouraging, because they suggest that all 4.2K applications involving RF SQUIDs are likely to prove to be viable at 77K also, though some may have to wait for good DC SQUIDs. Actually the limit on performance is often set by environmental factors and not by SQUID noise, and can be $10^{-3}\Phi o\text{-}Hz^{-1/2}$ or higher. It is clear that High T_C devices considerably short of the optimum would be adequate for such situations.

Note that the predicted noise limits are not fundamental. But SQUIDs with much lower noise would be physically too small to couple without very subtle double transformer coupling, which has never been achieved, even at 4.2K.

The prospects for SQUIDs if T_C's should reach 300K are still obscure. It is likely, though that they will be considerably superior, in noise and flexibility, to flux-gate magnetometers.

High T_C SQUIDS-progress

Both RF and DC SQUIDS have been demonstrated at temperatures up to 85K and higher using YBaCuO material.

The RF devices are bulk types and based on point contact Josephson junctions, and have now been fabricated in the quite sophicticated two hole geometry much used in 4.2K technology. As yet they are difficult to fabricate, and unstable against thermal cycling. The flux noise of the better ones, however, is now within a factor of 2 or 3 of the expected limit. Applications may not be far off-though caution is needed until 1/f performances have been reported.

The DC SQUIDS have included structures involving thin films. But good tunnel junctions have not been made yet, nor is it clear if they ever will be, because interlayer sharpness on a scale of the coherence length (1nm) will be needed. Nevertheless much work, worlwide, is being invested in trying to solve this problem. Meanwhile the active elements in all High T_C DC SQUIDs to data have been weak intergranular links: even with classical superconductors such granular devices are always characterised by large Johnson noise, strong temperature dependence, and severe 1/f noise and this is repeated in the noise performance (Table 1) so far achieved at high temperatures.

In summary, initial progress with high T_C SQUIDs is very encouraging. They may soon replace 4.2K SQUIDs in the least sensitive applications. Reaching the High T_C noise limit requires the development of a tunnel junction technology or some alternative. No progress has been made on this.

Magnetometers and gradiometers

Magnetic flux is usually applied to a SQUID by an input transformer connected in series to one or more superconduction pick-up coils. With a single pick-up coil, of diameter about 4cm (optimum matching inductance), a field of about $10^{-15}T\text{-}Hz^{-1/2}$ will result in flux of order $1 \times 10^{-4}\Phi o\text{-}Hz^{-1/2}$ being applied to the SQUID (unity S/N ratio for commercial RF SQUID). For DC SQUIDs with lower noise, greater field sensitivity is obviously possible. In practice one often accepts the opportunity to achieve the same sensitivity with a smaller pick-up coil such as can be accomodated on-chip.

Pick-ups are often configured gradiometrically. For example, coaxial coils of equal area located along a base line at z = 0 and 2D (1 coil at each site) and at z = D (2 coils wound opposite to z = 0, D ones) will be insentivite to uniform fileds B_z and uniform gradients dB_z/dz, and detect only the spatial differentials of the second order (d^2B_z/dz^2) and higher. Such gradiometers strongly select nearby (signal) sources at the expence of distant (unwanted) ones. They are excellent at rejecting the geomagnetic field and its fluctuations, and other spurious signals such as the ambient 50/60 Hz field. However precise balance is important, and can be difficult to achieve.

An approximate examples may be useful:

A dipole m located at z = -h below a gradiometer of order n and dimension D will produce an effective magnetometer signal (valid within a factor of 3):

$$B_{sig} = (m/h^3) \times (D/h)^n \times (1/2n) \text{ if } h >> D$$

or $$B_{sig} = (m/h^3) \times (1/2n) \text{ if } h < D$$

A $1m^3$ piece of soft iron placed in the earth's field (say $10^{-4}T$) yields, for an n = 2 gradiometer of baseline 30 cm:

$$B_{sig} = 10^{-4} \times 0.3^2 \times h^{-5} \times 1/2$$

If a Bsig = $10^{-15}T$, say, can just be detected, unity signal to noise ratio would be obtained for h = 85m.

It is important to note that a superconducting pick-up coil contributes no noise to the system. Thus even when noise dictates a 4.2K SQUID, pick-ups might be fabricated from high T_C material and run at 77K or higher. Simplicity of refrigeration and reduced stand-offs (with increased signals) are obvious advantages. It is important to recall, though that we are probably still further from having T_C wire that is suitable for such coils than we are from having practical high T_C SQUIDs!

Applications

General

At helium temperatures (4.2K and up to about 10K) SQUIDs have been applied to MAD, MADAIR, gravity gradiometers for navigation and mineral surveying, and to searches for gravity waves, magnetic monopoles and quarks. Work has also been done on contactless non-destructive testing of steel for cracks and stress, on ELF detection, on NMR, NQR and EPR, and on the susceptometry of micron sized specimens. (INTERNAL WAVES?) Commercial uses include gravimetry and magneto-telluric measurements of deep strata conductivity for mineral prospecting.

The most dramatic developments in SQUID technology are being fuelled, however, by interest in magnetophysiology and in particular by brain monitoring. Here there is a perceived need for up to 100 channels of thin film gradiometers for magnetic mapping above the skull. For such applications small pick-up coils are wanted: in turn that requires very low noise DC SQUIDs ($\Phi N = 2 \times 10^{-6}\Phi o\text{-}Hz^{-1/2}$). It seems likely that 4.2K will continue to be used for these and similar applications.

We now look at a number of these applications (MAD, gravity gradiometers, ELF antennas, NDT and magnetic resonance) which are of defence relevance. (INTERNAL WAVES?)

MAD

Programmes on SQUID MAD have been run for several years by ARE (Slough), in the US MAD and MADAIR have been extensively studied at NCSC Panama City, which published target tracking to 1'in 100' and field gradient results of $3 \times 10^{-14}T\text{-}Hz^{-1/2}$ over 10 years ago.

Attention is currently concentrated on so called 5-axis magnetometers using wire-wound coils on ceramic substrates. Coil balancing, using small movable superconducting tabs, is important and difficult.

The SQUIDs used are commercial RF ones, suggesting in the light of the noise discussion above that high T_C devices might prove adequate for MAD. But ARE is interested in having more sensitive (DC) SQUIDs for its current programme and also for a possible sonobuoy application: effective field sensitivities of $10^{-16}T$ in the range 5mHz to 500Hz have been mentioned. This implies SQUID noise sensitivities below $3 \times 10^{-5}\Phi o\text{-}Hz^{-1/2}$ and 1/f performance much better than is currently available even at 4.2K

The early use of high T_C SQUIDs does not seem very likely in this field.

Gravity gradiometers

Inertial navigation systems require continuous input of corrections representing the local gravitational field. This can be made available by recording the local gravity gradient and performing appropriate integrations. Precisions of about 1-10E ($1E = 10^{-9}ms^{-2}/m$) are required.

MOD has a programme in this field. An RF SQUID is used to sense the difference between the displacements of two spring-restrained proof masses separated by a distance of about 10 cm. Sensitivities of about 300E have been achieved, at which the system is SQUID noise limited with $\Phi_N = 1 \times 10^{-4}\Phi_0\text{-Hz}^{-1/2}$. Realising 1-10E will probably require both mechanical, thermal and SQUID noise improvements, which are in hand and will need 4.2K. For these and other reasons, early high T_C operation is not likely, though it cannot be ruled out long term.

Finally...no, gravity gradiometers cannot detect submarines! At a range of 500m, a mass of 10^4 tonnes produces a gradient of only 10^{-2}E. Worse, in a neutral buoyancy situation any effects would only be second-order: a spherically symmetric mass distribution would produce no signal at all.

ELF antennas

The NRL Washington programme worked on towed SQUID magnetometer antennas. The interest arose from a perceived need to detect signals centered on 80Hz with a bandwidth of 30Hz. at an attenuation level (due to skin depth effects) of 220dB relative to a 1 $V\text{-}m^{-1}\text{-}Hz^{-1/2}$ field strength at the surface. In magnetic field terms this indicates a sensitivity of $10^{-14}T\text{-}Hz^{-1/2}$, within the capability of RF SQUID systems.

To achieve such a sensitivity to ELF signals it was necessary to eliminate the motion-related noise caused by angular movement of the pickup coil with respect to the earth's field. This was done by maintaining an orientational stability of less than 4×10^{-6} radians for the towed structure (at 7 knots) which reduced the motion noise to $10^{-10}T\text{-}Hz^{-1/2}$, 86dB above the target value. This remaining noise was eliminated by an adaptive process based on combining the main output with the output from two other coils strictly orthogonal to themselves and to the main coil. The target performance was achieved and an ELF signal detected at 100m depth in sea water. The lack of further engineering development seems to be related to the failure to build a major ELF transmitting station, and not to inadequacies of the SQUID system.

Conventional RF SQUIDs were well able to deal with the signals envisaged, and the frequencies are well above the 1/f range. Thus high T_C SQUIDs use may be expected to feature strongly in any future programme. There would be considerable cryogenic simplificationsm.

NDT and magnetic resonance

SQUID magnetometers have been applied (NRL) to locating buried metal objects (such as pipelines) to which electrical contact can be made, and to the detection of gross flaws in them. Recently (MIT) they have been used to detect corrosion currents flowing in metal structures. Also SQUID gradiometers have been used (Strathclyde) to detect distortions which cracks and other flaws introduce into a polarising field provided by a permanent magnet; the magnet is carried in the same container as the SQUID. The MIT and Strathclyde techniques are contact free, and can 'see' through non-magnetic regions up to 5-10cm thick.

All of the techniques are limited by environmental, and not SQUID, noise. Indeed the last depends more on the continued periodic behaviour of the sensor in the presence of fields as big as 0.05T than on the fundamental sensitivity of the device. Thus all are strong candidates for the application of high T_C.

Of course, there may not be much direct MOD interest in NDT as such. But the polarisation method, which can detect small changes of magnetisation in the presence of a large field, may.

APPENDIX 8. 8B

APPLICATIONS OF SUPERCONDUCTORS IN ANTENNA SYSTEMS(*)

INTRODUCTION

The new ceramic 'high temperature' superconductors are potentially of use in antenna systems, assuming liguid nitrogen cooling is available. The following areas have been considered:

Electrically small antennas
a) at VHF and UHF
b) below 1MHz

Electrically large antennas
a) centimetric array feed systems
b) millimetric antanna feed systems

Coaxial cables
Cavity resonators
'SQUIDS' for ELF reception.

A detailed assessment of these possible applicatins has been made.

2. ELECTRICALLY SMALL ANTENNAS

Figure 1 shows small dipole and loop antennas. Their input impedance is

$$Z_{in} = R_{rad} + R_{loss} + j^{x} \qquad (1)$$

The antenna O for these antennas is given by

$$1/Q = \delta f/f_0 = (R_{rad} + R_{loss})/X \qquad (2)$$

where δf is the 3 dB bandwidth. The efficiency is

$$\eta = R_{rad}/(R_{rad} + R_{loss}) \qquad (3)$$

$$\text{Then} \; \eta \; \delta f = R_{rad}/X. \qquad (4)$$

Electrically small antennas are usually inefficient due to $R_{rad} < R_{loss}$. *A super-conducting antenna can have a much higher efficiency.* (Refs. 1,2) However, this is at the expense of bandwidth as the efficiency bandwidth product at an operating frequency fo is independent of the loss resistance, from equation (4).

Table 1

Frequency	dipole η	dipole (Hz)	loopη	loopδf(Hz)
1 MHz	0.006	10^{-3}		
10 MHz	0.17	0.3		
100 MHz	0.86	700	0.009	10^5

For 1/a and D/d fixed for the dipole and loop antennas of Fig.1, it can be shown (from formulae in ref. 3) that

$$\eta \; \delta f/f \; \alpha \; L^3 f^3 \text{ or } D^3 f^3. \qquad (5)$$

and that if $R_{loss} >> R_{rad}$,

$$\eta \ne \alpha \{L^2 f^{3/2} \qquad (6a)$$
$$\{D^4 f^{7/2} \qquad (6b)$$

Thus although the efficiency-bandwidth product behaves similarly with antenna size in wavelengths, the efficiency variation is markedly different. Numerical values have been worked out for a dipole with length 21 = 4 cm, and a loop with diameter D = 4cm, and are shown in table 1.

(Note that at 100 MHz, $\eta \; \delta f$ is 600 (dipole) or 900 (loop).) Thus an operating frequency at V/UHF, say 100-400 MHz, offers a usable bandwidth. The small dipole is already relatively efficient, but the small loop is not. *Superconductivity can give the small loop* similar efficiency and bandwidth to the small dipole.

A vertical small dipole gives an omni azimuth, vertically polarised, radiation pattern. A horizontal small loop gives an omni azimuth, horizontally polarised, pattern. A combined dipole *and loop could offer circular polarisation with an omni azimuth* pattern.

3. ATMOSPHERICALLY NOISE LIMITED CASES

At frequencies below about 1 MHz, electrically small antennas *are used for reception under "atmospherically noise limited"* conditions. Broadcast reception radios often contain a ferrite core multiturn loop antenna. Could superconductivity offer a smaller alternative?

For the ferrite core antenna.

$$R_{rad} \; \alpha \; N^2 \mu_{eff}^2 \qquad (7a)$$

$$L \; \alpha \; N \mu_{eff} \qquad (7b)$$

$$\text{whence, } \eta \; \delta f \; \alpha \; N \mu_{eff}. \qquad (8)$$

(N = no. of turns, μ_{eff} = effective permeability of the ferrite core).

For a loop or dipole with D or 21 = 4cm, at 1MHz, $\mu \; \delta f \sim 10^{-5}$

If $\mu_{eff} = 10^4$, N = 10, a ferrite core multiturn loop therefore has

$\eta \; \delta f \sim 1$.

Weeks (4) quotes a ferrite core antenna of 1" diameter and 10" long as having

$= 10^{-6}$ at 1MHz.

This then allows a bandwidth $\delta f \sim$ 1MHz, as required for broadcast reception.

Atmospheric noise at 1 MHz is about 70dB above room temperature thermal noise, so that efficiencies down to 10^{-6} still leave the atmospheric noise dominant (compared to $F.T._o$). The ferrite core antenna is therefore very well matched to its application. Superconductivity allows a choice of efficiency within the constraints $\mu \; \delta f$ constant, which in this case offers no improvement.

4. MICROSTRIP FEED NETWORKS

James, Hall and Wood[5] suggest that the dissipative losses associated with microstrip lines are one of the major limitations with microstrip antennas. They also point out that conductor loss is considerably greater, than dielectric substrate loss at microwave frequencies. Figure 2 compares conductor and dielectric losses for various cases. This suggests that a factor of 10 reduction in loss (in dB per unit length) could be achieved with superconducting lines.

(*) Written by M.S Smith of STC Technology Ltd., UK

For planar antenna arrays fed with a microstrip network, including an end fed travelling wave array, the feeder losses limit the achievable gain. As the array size increases, the directivity increases, but the feed losses go up and eventually overtake the increased directivity effect, as shown in figure 3 (taken from ref. 5). In ref. 5 it is shown that

$$G = G_{el} + 10 \log_{10} 2L^2/\lambda_0^2 - 3L/2 \,.\, F \qquad (9)$$

where F is the feeder loss in dB/unit length and L is the (square) array dimension.

The maximum gain is where

$$L = 5.776/F$$

An example with frequency 12 GHz, F = 0.075 dB/cm gives L_{max} = 77cm and G_{max} - 30dB.

If F were reduced by a factor of 10, G_{max} -~50dB, a very significant improvement.

5. MILLIMETRIC ANTENNA FEED SYSTEMS

At millimetric frequencies, standard size waveguides used to feed say a reflector antenna are very lossy (6). Two ways round this problem are currently available, (i) oversize waveguides, and (ii) beam waveguides.

For oversize waveguides, with the operating frequency for above the cut-off frequency of the dominant mode, the guide attenuation is considerably reduced. However, such guides are prone to spurious mode generation, so that stringent tolerance requirements become necessary. Aperture blockage may also be a problem with some reflector arrangements.

Beam waveguide systems, as shown in figure 4, can be used with suitably configured reflector antennas, with low loss. However, the cross-sectional dimensions of a beam waveguide are necessarily much larger than standard size waveguides, and are comparable with the dimensions of oversize waveguides with low loss.

Superconducting millimetric waveguides would allow the use of standard size waveguide feeders, and all the usual microwave reflector antenna configurations would then be available.

6. CABLES, RESONATORS AND SQUIDS

The use of coaxal cables is limited to lower frequencies by conductor loss. Refs. 7 and 8 reported measurements on superconducting coaxial cables at liquid helium temperatures, which demonstrated low losses up to 12 GHz. (For example ~ 1dB/km at 1 GHz). With superconductors now available at liquid nitrogen temperatures, their findings become that much more practical.

At very high frequencies (microwave and mm wave), loss in the walls of cavity resonators prevents high "Q" from being realised. Superconducting cavities offer much higher Q. Refs. 9-12 describe work on superconducting resonators at liquid helium temperatures.

Q's of 10^5 to 10^{11} for frequencies between 1 MHz and 35 GHz were described. Again liquid nitrogen colling makes these applications more practical.

"SQUIDS" can be used as antennas for ELF reception (refs. 13-15). The sensitivity required for ELF reception is very close to the limit for devices at liquid helium temperatures. A liquid nitrogen temperature SQUID would be cheaper to use, but the RF front-end is then at 77K rather than 4.2K. This makes the liquid nitrogen SQUID less sensitive (ref.16), and it will be very difficult to achieve the required sensitivity. Further developments should demonstrate what sensitivities are achievable.

7. CONCLUSIONS

The table below summarises the potential of each possible application examined:

Application		Superconductors advantageous?
Electrically small antennas		
a)	VHF/UHF	√
b)	below 1MHz	X
Electrically large antennas		
a)	centimetric array (distribution network)	√
b)	millimetric reflector (feeds)	√
	Coaxial cables	√
	Cavity resonators	√
	'SQUIDS' for ELF reception	?

8. CURRENT WORK

Materials research and development is being undertaken at STL, to provide a screen printing process to make superconducting devices. The first measurements will be of a microstrip transmission line, over a wide range of frequencies. A loop antenna with a tuning and matching section (also superconducting) is also planned.

REFERENCES

[1] Walker, G.B. and Haden, C.R. "Superconducting antennas", J. Appl. Phys., 40, pp. 2035-2039, 1969.

[2] Khamas, S.K., Mehlen, M.J. Maclean, T.S.M. and Gough, C.E. "High T_c superconducting short dipole antenna, Electronics Letters, 24,pp. 460-461., 1988.

[3] Jasik, H. "Antennas Engineering handbook", McGraw-Hill, 1961.

[4] Weeks, "Antenna Engineering", McGraw-Hill, 1968.

[5] James, J.R., Hall, P.S. and Wood, C. "Microstrip Antenna theory and Design", Peter Peregrinus, 1981.

[6] Bhartia, P. and Bahl, I.J. "Millimeter Wave Engineering and Applications", Wiley, 1984.

[7] Nahman, N.S., "Miniature superconductive coaxial transmission lines", Proc. IEEE, 61, pp. 76-79, 1973.

[8] Chiba, N.Kashiwayanagi, Y and Mikoshiba, K. "Frequency characterictics of superconductive coaxial lines", Proc. IEEE, 1973.

[9] Harting, W.H. "Superconducting resonators and devices", IEEE, 61, pp. 58-70, 1973.

[10] Danielsen, M, "Superconducting lead cavities at 35 GHz", Proc. IEEE, 61, pp. 71-75, 1973.

[11] Vig, J.R., and Gikow, E. "A superconductive tuner with broad tuning range", Proc. IEEE, 61, pp. 122-123, 1973.

[12] Jimenez, J.J., Benard, J, Sudraud, P and Septier, A. "Attainment of a low noise high power and highly stable Gunn oscillator by coupling to a superconducting cavity", Proc. IEEE, 61, pp. 123-124, 1973.

[13] Burrows, "ELF communications antennas", Peter Peregrinus, 1978.

[14] Clarke, J. "Low frequency applications of superconducting quantum interference devices". Proc. IEEE, 1973.

[15] Wolf, Davis and Nisenoff, "Superconducting ELF magnetic field sensors for submarine communications", IEEE Trans. COM-22, 1974.

[16] Pegrum, C.M. and Donaldson, G.B. "High critical temperature SQUIDs how sensitive will they be".

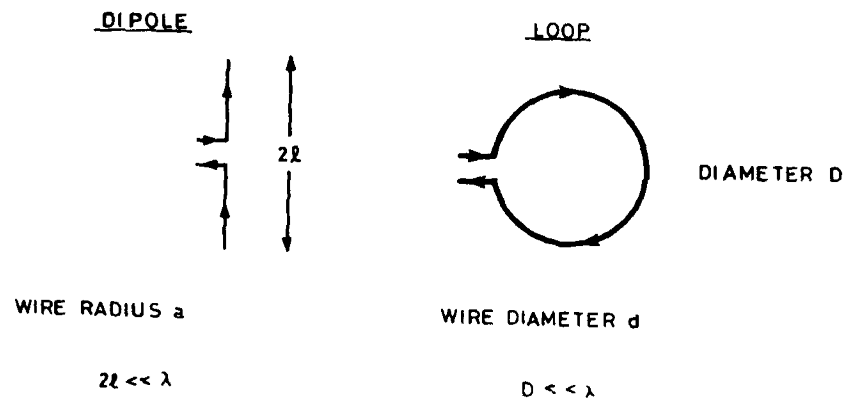


Fig.1 Dipole and loop antennas

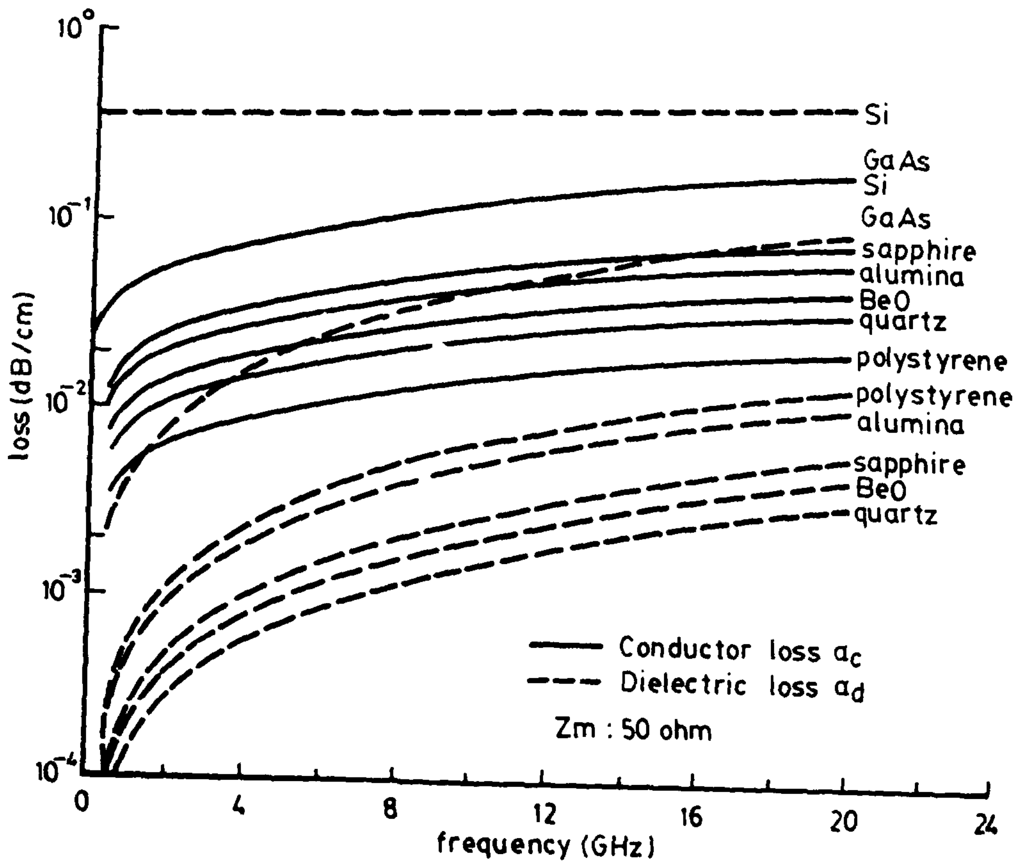


Fig.2 Conductor and dielectric losses as functions of frequency for microstrip lines on various substrates (Reproduced from Gupta et al. (1979))
h = 0.635 mm except for GaAS and Si where h = 0.254 mm

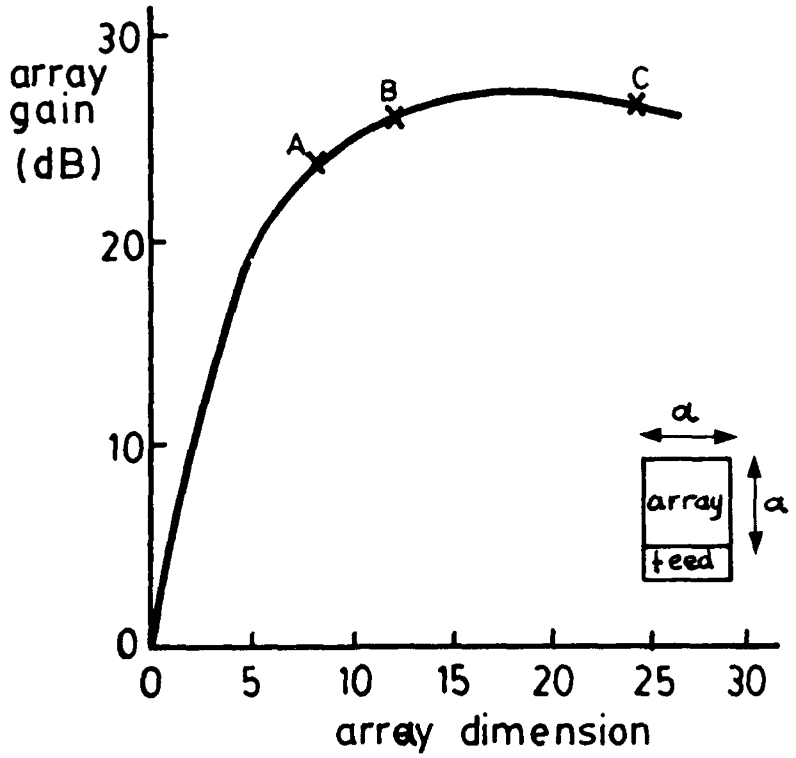


Fig.3 Gain limit for a microstrip end-fed travelling-wave two-dimensional array A, B measured values, C estimate based on 0—1 dB/λ_m

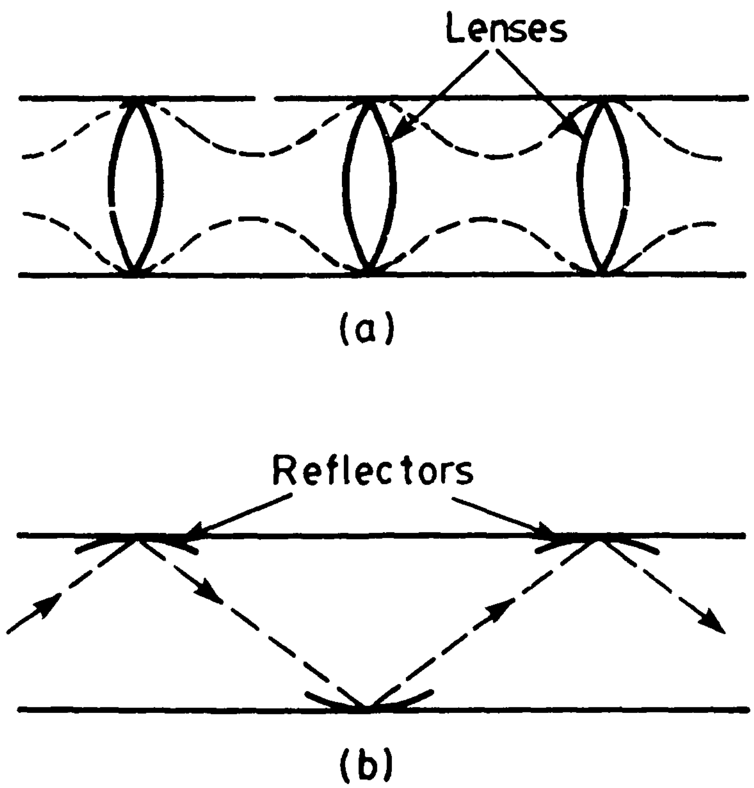


Fig.4 Beam waveguides using (a) lenses and (b) reflectors

8.9 ARTIFICAL INTELLIGENCE AND NEURAL NETWORKS

8.9.1 Introduction

This paper considers the possible role of Artificial Intelligence techniques, in future military satellite communication systems, into the next century. It is a speculative activity; firstly because generic military SATCOMs systems will experience considerable changes over the next few decades as new technologies emerge and mature; and secondly because AI is still a relatively new discipline, with its future role and potential not clearly defined. It is however probably safe to assume that future milsatcom systems will become increasingly complex, support a greater diversity of users, and incorporate sophisticated anti-threat capabilities. And that a particular outcome of these changes will be a considerable growth in the system management and control overhead.

Existing SATCOM systems incorporate computer support for tasks involving the processing of data from various information sources, which have to be executed in real time. Examples include; telemetry telecommand and control, and the management of multiple access schemes. The demand for this types of function will increase, but with greater power and autonomy, will follow the predicted increase in system complexity, and user growth and diversity. The question is whether this can be implemented by means of conventional data processing techniques or will it be necessary to employ more novel types of solution such as AI. Another important question which needs to be addressed within our general area of consideration is whether intelligent processes can be employed in the system whilst still maintaining the necessary degree of system predictability. True intelligence is characteristically undeterministic which may not be tolerable.

In addressing these questions it was considered necessary to first say what is meant by AI, and what in general its scope and potential areas of application are. This is intentionallly a selective view, as the field of AI is a broad area of activity, and not easily summarised. Also certain areas within this field may be said to offer more obvious applications to militay SATCOMs than others. Consideration was then given to identifying areas of military SATCOM systems where the realisation of future system capabilities may require that unconventional solutions such as AI are considered. Finally an attempt was made to combine the two perspectives, and say in which case AI may provide an appropriate solution.

8.9.2 Artificial Intelligence an Overview

AI techniques are appropriatefor particular classes of problem which in general do not respond to standard programming methods. The two approaches should therefore not be regarded being in any way interchangeable. The term AI may be used to describe everything within the area of non-procedural, non-algodithnic and partially non deterministic problem solving techniques.

AI may for our purposes be subdivided into three main areas of activity:

- Knowledge Based Systems,
- Neural Networks,
- Conventional AI Techniques,

8.9.2.1 Knowledge Based Systems

Knowledge based systems are being researched and developed to deal with classes of problem which resist a sutructured approach to achieving a solution. The principle differences between knowledge based systems and conventional computer programs are characterised below.

Data Processing	Knowledge Engeneering
Representation and use of data	Representation and use of knowledge
Algorithmic	Heuristic
Repetitive process	Inferential process
Manipulation of data bases	Manipulation of knowledge bases

A knowledge based systems consists of one or more pools of knowledge, termed domain knowledge. Within this domainknowledge, is encoded the expertise of one or several human experts, in the form of condition-action pairs. If a defined condition occurs, a related action is triggered and executed. This can of course lead to the treggering of another condition-action pair. Conditions can consist of single or compound statements, actions can consist of one or several actions. The knowledge in a KBS is said to be coherent if no contradictory or ambiguous condition action pairs exist. A query (question) is given to the system which will then attempt to find an answer in its knowledge space with respect to given constraints. Systems of this type are appropriate for application to problem domains where knowledge (empirically acquired experience) is the only economical or technically feasible method of solution.

8.9.2.2 Neural Networks

Neural networks, as a technique for data processing, have gained considerable attention over the last few years. Some aspects of the connectionist philosophy are well understood and already established in certaion areas of science; examples being Boltzman machires and adaptive systems. More recently the massive parallel approach to solving problems has gained some prominence. The techniques in this case are deriveo from biological models of signal propagation in nerve fibres Neural Networks aer generally used to associate new data with data acquired through training sessions. During training sessions, the acquired data is classified and categodised using one of several learning techniques. The undisputed advantage of neural processing is the speed with which the nets are able to map incoming data into groups or patterns generated during training sessions. It is not necessary for this input data to exactly match the established data groups or patterns, as a considerable range of deviation can be tolerated. So far most neural processing is carried out with general purpose computers using programs designed to simulated the behaviour of neural nets. Some companies and universities are developing special purpose hardware that will speed up program execution time by a factor between 10 and 1000, and resemble more closely the operation of actual neural networks.

8.9.2.3 Conventional AI Methods

Conventional AI methods include those techniques, that use programming languages such as LISP and PROLOG to arrive at decisions to problems. These languages are generally used to apply a decision process (or at least provide decision support) with respect to abstract data from sensors or other sources. Strategieb or tactics can be contained implicitly in these programs to achieve a coherent line of decisions. The advantage of LISP and PROLOG is that their powerful semantic and syntax permits complex data structure definition. They are generally installed and run on conventional computers although some dedicated hardware support exist in the case of LISP

8.9.3 Possible Candidate Areas of Future Military SATCOM System

The following were considered to be representative areas of future growth where successful implementation may call for novel solutions, either advances in conventional processing techniques or application of the AI methods described above. These are not intended to serve as a comprehensive set of all potential recipients of AI. It is however considered reasonable to use these specific cases as a basis for drawing more general conclusions.

. Link Performance Management,
. Satellite Access,
. Electronic Counter Counter-measures.

These were evaluated as follows:

8.9.3.1 Link Performance Management

The performance and capacity of any satellite communications system is largely dependent upon the link budget. It is generally downlink EIRP which is at a premium, and for each small terminal user there is a potential trade off between power, antenna size, and achievable data rate.

Suitable margins are included in link equations to provide for the effects of propagation loss, antenna pointing, and engineeding implementation. These will depend upon the frequency band, area of operation (including the weather and the elevation angle to the satellite). In general military systems are not as controllable or as predictable as their civil counterpart and therefore large link margins are allocated. This of course reflects teh worst case. Typical link margins at EHF for moderate elevation angles might be ; 8dB for 99 availability at 20 GHz, and 20dB for 99 % availability 44 GHz. Actual link ferformace will vary considerably depending on detailed implementation and the effects of weather

In view of the size of link margin required for future SATCOM systems operating in teh EHF frequency bands, a link management capability which minimised the required link margin in response to the prevailing environmental conditions would offer both LPI and power saving benefits

The link management function to maintain optimum down link power would have to take account amongot other things of environmental factors, the effects of jamming and user capability. This would be an extremely complex tast because of the number and unpredictable nature of the parameters to be considered. and an extremely critical onve because error would result in loss of the link. It may be taht implementation of tis is beyond the capabilities of conventional algorithmic and procedural programming methods. However the fact that the availability of the link would, in the case of an AI solution, be undeterminictic to some degree tends to rule out this type of solution, even if it were otherwise appropriate

8.9.3.2 Satellite Access

Milsatcom transponders may be required to simultaneously handle a large number of links, using different forms of traffic and protocols, and involving a number of terminals. It is also characteristic of the military oxfor accesses to come and go unpredictably, and requirements to vary rapidly. The system needs to be able to respond to this, and in periods of high demand also take account of user priority.

The multiple access problem is that of allocating and implementing the sharing of transponder capacity between a number of terminals, most of whom are operating on a single channel per carrier basis. Individual users may access a link by prior arrangement or in conjunction with a polling or request channel. A group of users (a net) may employ their own protocol over an allocated channel by pre arranged assignments of by random access contention.

It is clear that for the system to allocate its resources most effectively more control must reside in the space segment. Thus making effective use of the satellites direct access to a whole range of user parameters. This in turn would reduce the amound of traffic between the ground and space segment required for command and control purposes. It is however felt that the level of on-board control appropriate to the satellite could be met by algorithmic and procedural programming methods. This is not to say that kbeyond some level of autonomous operation intelligent processes would not offer some advantage. The criteria in this case however must be. how much autonomy can be given to user access functions within a SATCOM system, in particular the space segment, whilst still maintaining a satisfactory level of overall control.

It is considered in this regard that, whilst some control should reside within the satellite, this should be limited to what can be achieved by the application or conventional processing methods. With the application of intelligent processes, should they be found to offer some distinct advantage in the task of overalll user access control, being restricted to the ground segment. In this way a more secure and accessible interface would exists between the processes and overall system control. It is however not considered that intelligent processes would offer any advantages for this application, which were not outweighed by the stated disadvantages.

8.9.3.3 Electronic Counter- Measures (Antenna Nulling)

In order to maximise communications capacity under jamming it is necessary to remove as much jammer power as possible by techniques on board the satellite and/ or at ground terminals. While Anti-jamming protection may usefully be applied at terminals, it may also be worthwhile to prevent and uplink jammer from capturing the satellite transponder power. One powerful AI techniques which will be further developed for future satellite application is antenna nulling

Antenna techniques on board the satellite may help to alleviate the effects of uplink jamming. A reduced uplink coverage may be used to enhance the wanted signal at teh expense of a jammer provided the two are physicallly qelll separated. This is a simple and obvious measure but required a large aperture antenna and provide only limited discrimination it also conflicts with the requirement for global coverage.

The concept may however be extended to the provision of an array of spot beam antenna, with selection of the appropriate coverage region. This would be integrated as a Multiple Beam Antenna (MBA), with a number of feeds sharing a common dish reflector or else weveguide lens structure. An MBA has the advantage of providing flexible coverage with high gain but in its simplest form gives only limited jammer rejection.

Improving jammer rejection of specific interference sources may be achieved by combining the signal from two or more elements of an MBA

The use of several antenna elements together with both phase and amplitude control and combination, may permit considerable flexibility as a nulling antanna, and may allow simultaneous nulling of several interference sources. It might also be possible to synthesize area nulls, for example over hostile territory without knowledge of specific jammer location

Such sophisticated antennas would represent considerable complexity on a spacecraft. Efficient control is a major aspect and this may be either by remote telecommand or locally through on-board adaptive means if the latted can be achieved. It is considered that despite the possible complexity of this control task the need for precision requires a determinictic

solution of the sort offered by conventional rather than AI methods.

8.9.3.4 Satellite Management

A considerable amound of resources in the ground segment are currently dedicated to satellite management tasks. These can be divided into three main categories:

- In-orbit control
- Status and health monitoring
- Communication management

The satellite management workload in future systems will undoubtedly increase, particularly in the latter two categories. The predicted trend in future systems for increased complexity and autonomy, cna only be achieved by the incorporation of more extensive monitoring and control facilities within the system. Using a convential TT&C arrangement ie the detailed monitoring and control of the satellite being exercised from the ground, a pro-rata increase in telemetric traffic will result. If however a level of decision and control could be delegated to the space segment, so that only top level command and reporting between the ground and space segments were required, the volume of telemetric traffic could be proportionally reduced.x

The majority of low level satellite management tasks are inherently routine and predictable and are therfore suitable for implementation using conventional procedural programming methods. For contingencies taht cannot be easily programmed in a procedural manner, an expert system may be an appropriate solution.

8.9.4 Conclusion

It can be stated generally, that a system which satisfies a given requirement should not be modified if this modification offers only marginal improvement. This is especially true with respect to AI. Therefore any application that is realisable using algorithmic and procedural programming methods should be implemented in this way. Most cases for which a heterogeneous solution space exists, can be realised as rule based systems.

The general constraint which applies when considering the use of intelligent processing in systems is one of predictability. Thus a system that includes intelligent processing is dependent upon an array of various senson inputs to arrive at a solution. This leads to the problem that decisions arrived at by the system cannot be accurately predicted but only traced back to the starting statet of the specific computation. This fact would make the SATCOM systems, or at least those parts which were the subject of the decision, undeterministic to some degree, a feature which cannot be tolerated in this sort of application. This is compounded by the accesssibility of the satellite. Any error which has an impact for example on the pointing angle of the antennas, would make the satellite completely unusable.

It can be concluded that most of the control and data processing tasks of a satellite communications system are primarily of a procedural and/ or algorithmic nature. It is therefore predicted that the solution to such tasks will continue to be solved by conventional programming methods. For some of the satellite management tasks, with a variety of possible actions, a small expert system might be appropriate and the only clear example of this.

8.10 ROBOTICS AND CONTROL

8.10.1 Introduction

The term robotics is widely used to describe methods of executing functions or operations, either fully autonomously or with the minimum involvement of man. This is needed where the operation is too difficult, too hazardous, too uneconomic, or simply impossible for man to carry out. These circumstances frequently occur in space systems. Perhaps the most common example of robotics is the autonomous operations employed in satellite attitude and orbit control. Other robotic roles now beginning to be employed are the mechanical handling aids used for spacecraft recovery and repair and the rendezvous and docking systems to be used for coupling together several spacecraft.

8.10.2 Spacecraft Autonomy

Some degree of spacecraft functional autonomy has always been required, if only to ensure survival in the event of certain types of failure. These are commonly referrred to as fail safe modes and are intended to preserve essential services such as solar array power until such time as human intervention can occur and perhaps resolve the problem and take corrective measures. For geostationary satellites telemetry, command and control (TCC) communications can be continuous but for satellites in inclined eccentric orbits such as Molniya TCC communications are limited to periods when the satellite is visible to a ground control station. Systems which use such orbits therefore require a longer period of satellite autonomy in the event of failure.

In jamming or nuclear burst conditions, ground to satellite TCC communications may not be possible, perhaps when it is needed most. In these cases periods of autonomy of somewhat indeterminate elength will be needed before ground control can be re-established. Similarly ground control stations are themselves subject to failure and are targets for enemy attack. In such circumstances spacecraft autonomous operation would then become very desirable if not essential.

The following aspects of spacecraft operations can benefit from significant levels of autonomy:

a) Fault diagnosis and repair
b) Real time adaptation to traffic density
c) Real time adaptation to jamming conditions
d) Spacecraft houseteeping
e) Rendezvous and docking
f) Inter satellite links

In the 2000 to 2030 time frame more extensive autonomous operations are expected to be technologically possible and with consequent extended satellite life times and reduced ground support costs would be more affordable.

8.10.3 Autonomous Operations

8.10.3.1 Fault Diagnosis and Repair

Fault diagnosis in current satellites is carried out by using a limited set of telemetry data and associated pre-determined fault finding routines. Repair is then been implemented by a changeover, to a redundant subsystem. When these data and routines fail to provide a definitive diagnosis, analysis of all available information (including historical data) is carried out by engineers familiar with the system. As a result of this analysis further tests are carried out on the satellite using irregular configurations and non-normal conditions to obtain more information. This process may take several weeks in extreme cases and unless there is an in-orbit spare satellite, the system may be down for some time. It may, however, be possible under inter-operability agreements to obtain some communications capacity which can be used until a satellite is put in place.

In the future, the diagnosis process is at first, expected to evolve, to use both on-board and on ground expert systems. These will encapsulate the system design knowledge and failure modes and effects analysis which will have been carried out during the development phase. The expert system data base will continue to be updated and added to during the operational period by the uplink to the satellite computer's of new knowledge. This first step will allow decisions to be taken on-board the spacecraft in response to failures whose mechanisms which are already known or become known from operational experience. Failures outside the capability of the on-board expert system to resolve will still need to be referred to engineers on the ground. The consequences of this first step will be to extend the interval during which the spacecraft can operate autonomously and also reduce the staffing levels needed at the ground control stations. However, access to engineers with detailed knowledge will still be required to identify and provide solutions for those failures which are not predictable and for which there is no previous experience.

The next step may be to employ artifical intelligence (AI) methods, which will not only use an expert knowledge database but use learning processes to add further information. The learning system may use Neural networks in conjunction with 'What if' trials to resolve the very large number of possible combinations and permutations of failure modes and mechanisms. Parameter trend information derived from historical records stored on board together with data relevant at the time of failure is used to search the knowledge database for a failure which best matches the status information. Having identified the failure the knowledge database is searched for a reconfiguration of the system which will most colsely achieve the goal of maintaining a preset mission capability including satellite safety before implementing the changes. The mission requirements may change during the life of the SATCOM system and a new definition of it would then be updated from the ground control station or in principle, from any SATCOM terminal as the need arose.

Repair of the system would, as now, be implemented by reconfiguration to use redundant elements internal to the satellite if available or with in-orbit spare capability in other satellites of the system.

8.10.3.2 Adaptation to Traffic Conditions

A communications payload of the type described in section 12.2 which uses phased array antennas, distributed transmit and receive modules and on board signal processing, is capable of being reconfigured to operate in many different modes suitable to a variety of traffic conditions including jamming and propagation conditions.

The parameters which can be altered to suit the traffic requirements include:

Gain adjustment, beam shaping and steering and null positioning using antennas.
Reallocation of transmitter power and frequency band for up and downlink. Alteration of processing gain by algorithm choice. For systems with spatially distributed satellites e.g. clusters spatially, time and frequency division multiplexing of traffic using

inter-satellite links

A system sim... concept to the one for fault diagnosis and repair could ... implemented to autonomously manage and optimise traffic performance. A knowledge database of possible traffic conditions and corresponding optimum spacecraft payload configurations would be initially established during development and then updated progressively during the system life time by uplinking new algorithms.

8.10.3.3 Satellite Housekeeping

This is the area where autonomy is currently implemented to a high degree. The mechanisms however, are simple in concept and are implemented by closed loop servomechanisms. These range from attitude and orbit control systems using a variety of attitude sensor types appropriate to the class of orbit together with thrusters to compensate positional changes due to spacecraft drift and planetary influences, to Power system management and thermal control of the spacecraft on board environment. In future satellites these functions would be part of the task of the same on-board fault tolerant computers already identified for Fault diagnosis, repair and adaptation to traffic conditions.

8.10.4 The Autonomous Spacecraft Manager

The achieve the foregoing degree of autonomy the spacecraft must be designed to carry out many of the tasks which are usually carried out by the ground facilities. These would include

a) Health Monitoring
b) Performance Monitoring
c) Station and Attitude Keeping
d) Re-configuration to maintain health and performance
e) Re-configuration to optimise traffic handling
f) Keep a historical record of status, health previous failures and repair processes implemented

Designing a spacecraft to meet these requirements will require new architectures and technologies. It is likely that the ultimate system design will have a hierarchy such that dangerous conflicts will be avoided. Each subsystem and payload element will be required to have higher levels of self intelligence but with overall responsibility held by an On-board Spacecraft Manager. In general, a pre-determined plan will be used to control the operational activities, updated as required from the ground or automatically in orbit by monitoring user traffic and environment. Figure 8.10.1 below shows a possible architecture type which would perform these functions.

8.10.5 Rendezvous and Docking

Some of the future space system architectures envisage construction in orbit, e.g. the US Space Station Freedom. At the space station low earth orbit, it is practical for man to play a significant role in the construction and the 'Shuttle' docking process with the station. At higher orbits however - geostationary or highly elliptical orbits-the radiation environment and transport costs of a manned servicing vehice prohibit the use of man. As a consequence construction and repair processes will need to be carried out robotically. These processes will involve rendezvous and docking with the system under construction. If repair of the system or upgrade is also envisaged, similar processes will be required to remove the defective element and replace it. All such operations will require a very high degree of autonomy and precise spatial control at long range.

Provided the cost of these robotic systems and operations remain lower than the complete replacement cost of the satellite involved, this will become the mechanism employed to extend the system life time. In such a case refuelling of the operational spacecraft becomes essential and refuelling may then be accomplished by a tanker spacecraft docking with the operational satellite. The empty tanker spacecraft most probably being subsequently ejected from the operational orbit into deep space, rather than re-entry to low altitude recovery orbit.

There appears to be no technological reason why these robotic processes cannot be employed in the post-2000 era.

8.10.6 Inter- Satellite Links

The formation of optical or microwave links between spacecraft will require the use of automatic acquisition and tracking processes between the cooperating spacecraft. The pr... is similar to that for the detection of target aircraft by a missile except that the distance between the two spacecraft does not decrease with time. The target spacecraft is however cooperating and provides a good signal to acquire and track.

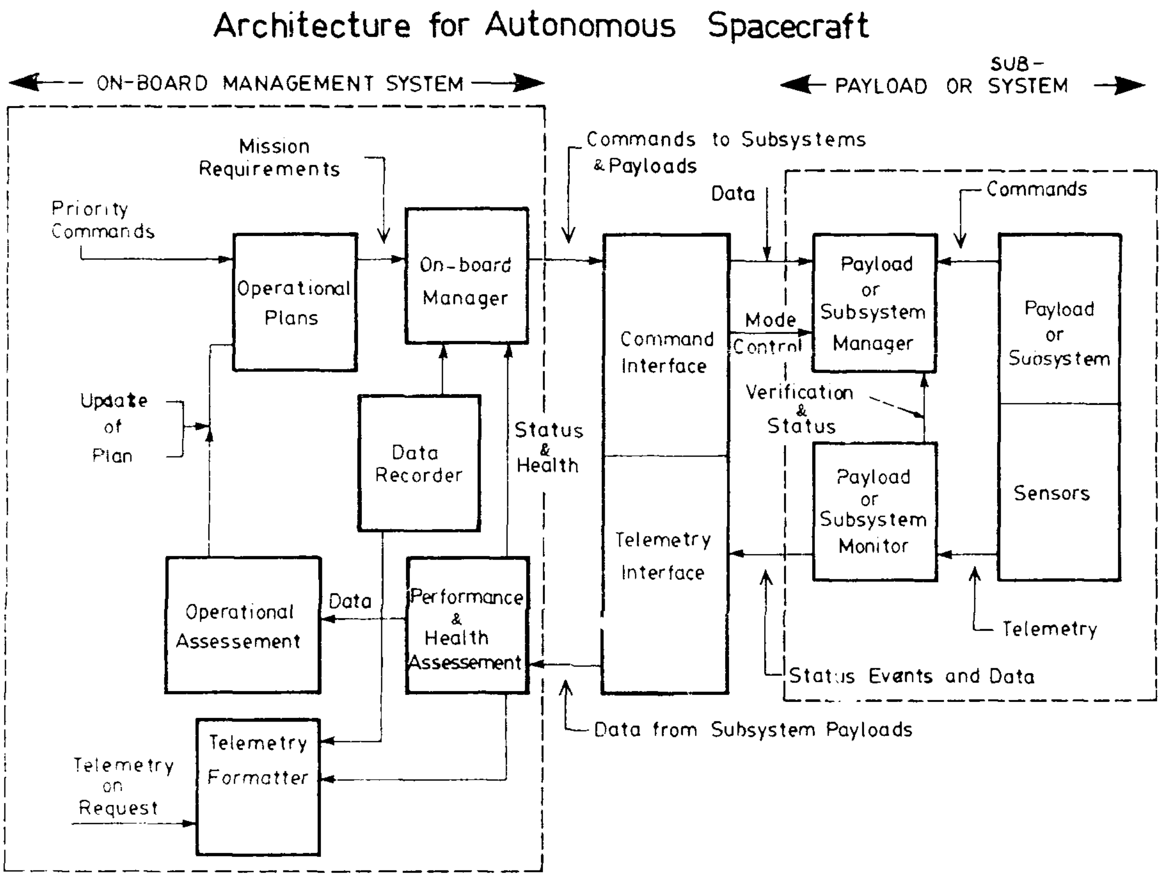


Fig 8.10.1 Spacecraft autonomous management system

8.11 POWER GENERATION IN SPACE

8.11.1 Introduction

Power systems for space use cover two main functions-Generation and Storage.

8.11.2 Power Generation

Power generation in space is produced by either photo-voltaic or thermal to electric conversion means.

8.11.2.1 Photo-Voltaic

This method is typified by the failiar solar array which consists of a deployable structure housing many individual photovoltaic cells. There is no theoretical limit to the size of the array except that posed by accommodating the arrays within the launcher volume constraints and the practicality of deploying or assembling them in orbit. Currently single arrays of about 10 Kw are feasible, but multiple arrays can obviously increase the generating capacity.

a) Photovoltaic Materials.

Table 8.11.1 Photovoltaic Materals and Energy Gaps

PHOTOVOLTAIC MATERIALS	ENERGY GAPS eV
Cadmium Sulphide (CdS)	2.4
Gallium PhosPhide (GaP)	2.2
Cadmium Selenide (CdSe)	1.7
Cadmium Telluride (CdTe)	1.55
Tungsten Diselenide (WSe_2)	1.5
Gallium Arsenide (GaAs)	1.43
Indium Phosphide (InP)	1.34
Zinc Phosphide (Zn_3P_3)	1.3
Copper Sulphide (Cu_2S)	1.2
Silicon (Si)	1.1
Iron Sulphide (FeS_2)	0.95
Germanium (Ge)	0.7
Lead Sulphide (PbS)	0.37
Lead Telluride (PbTe)	0.29
Lead Selenide (PbSe)	0.26

Table 8.11.1 shows the range of materials and their energy gaps and Figure 8.11.1 shows the theoretical efficiencies which can be obtained against Eneregy Gap and temperature. Cell performance using these materials is also influenced by their spectral response and degradation due to the radiator environment of the operating orbit.

Solar cells used in satellites to date have almost exclusively been made of Silicon which has a theoretical efficiency at 25 degrees C of 20%. Practical efficiencies currently achievable are 12 to 14%. Cells using Gallium Arsenide and Indium Phosphide are under development, and practical efficiencies achieved so far are 16 - 19% for GaAs and around 16% for InP. In the time frame 1990 to 2030 the efficiencies for each type may be expected to move towards the theoretical figures in Figure 8.11.1

GaAs and InP cells are more expensive to produce than Si cells but their greater efficiency and lower rates of degradation allow arrays of lower mass to be made. This in turn reduces launch costs and more than likely offset the cell costs. Nevertheless continuing research into other materials may yield further cost reductions. A particularly promising material for research is Iron Sulphide which is a frequently occurring natural material. It has a quantum efficiency of at least 33 % with visible light. Achievable efficiencies to data are very low 1 - 2%, but with research could be 15 20%.

b) Arrays.

i) Planar Arrays.

Solar arrays of this type consist of cells laid side by side on a substantially flat surface which can be either flexible or rigid. The surface is supported by a structure which can be folded or rolled up for the launch phase and subsequently deployed in orbit. This form of array is used on 3 • axis stabilised spacecraft as wing like structures which are ratated towards the sun so as to maintain the best incidence angle during orbit and seasons. For spin stabilised satellites the surface is wrapped around the spacecraft to form a facetted cylinder. In this case the satellite is itself steered so as to point the array towards the sun.

Table 8.11.2 shows typical current array masses and power outputs/unit area for silicon and GaAs cells.

This table shows that for an output of 5 Kw to be available from a flexible array at the end of life a silicon array would weigh 102 Kg against 78 Kg for an AlGaAs array i.e. 31% more. For a rigid array the silicon array would weigh 24% more.

Table 8.11.3 shows the characteristics which are expected to be available some 10 years hence. Significant improvements are expected for both Silicon and Gallium Arsenide cells. For the same 5 Kw at the end of life the mass of a flexible Silicon array would be 56 Kg against 39 Kg i.e. 43.5% more and for a rigid array the Silicon array would weigh 27 % more. These represent array masses for both material types of about 50% of current values.

GaAs cells retain their relative advantage over Si cells due to their inherent lower radiation degradation rate.

ii) Concentrator Arrays.

This type of array aims to increase the effective sun illumination falling on the individual cell by concentrating the light by diffraction, refraction, or reflection. The most suitable method for space use is by reflection and most development has taken place on a Cassegrain miniature concentrator.

Table 8.11.4 shows the characteristics for future concentrator array generating 5 Kw. The mass to power ratio for concentrator arrays are worse than planar arrays by factors of between 4.1 and 1.36:1 depending on comparison with flexible or rigid arrays.

However, concentrator arrays have an advantage over planar arrays is being less vulnerable to laser attack and have improved radiation performance. This arise because the concentrator limits the incident illumination on the active material to narrow angle. The improvement in radiation hardening can lie between 2:1 and 5:1 for AlGaAs and InP material cells respectively.

A second mechanism for improved radiation performance is by repair of the damage by annealing the cell. InP cells show promise of annealment at close to room temperature i.e. 20°C

(*) This Section was prepared with the assistance of the Royal Aerospace Establishment, Farnborough, UK.

which could be accomplished on board the spacecraft continuously. This could effectively eliminate radiation damage with time completely.

8.11.2.2 Thermal To Electric Conversion.

Such systems consists of a high temperature energy source, a heat transport system to Heat to Electricity convertor, and a heat transport system to a cold sink. Figure 8.11.2 shows a simple diagram of the system.

a) High Temperature Heat Source

i) Solar Dynamic.

In this system infra-red solar radiation is concentrated optically and the energy is either stored in a large thermal mass to eliminate the need for electro-chemical storage in batteries or is transported directly to the converter by for example heat pipes.

ii) Radio Isotope.

The heat source in this system is a radio isotope such as Plutonium 238 which has a long life. The heat source is tightly coupled to the thermoelectric convertor such as Silicon-Germanium thermocouples. An example of such a system is given in Ref. 1. This system requires no energy storage system and its long life time makes it suitable for deep space missions and orbits with many eclipses such as polar. It is inherently immune to radiation damage. The power available is only limited by cost and launcher mass volume constraints. A typical Radioisotope Thermoelectric Generator (RTG) which produces 250 watts for a period of 5 years has a volume of 0.2 cubic metres and a mass of 56 Kg. It is evident that to provide 5 Kw of power the mass would be 1120 Kg i.e. about an order of magnitude more than a current solar array.

iii) Nuclear Reactor.

This system is similar to the RTG in that heat is generated by nuclear reactions. They are only economic for power generation greater than 10 Kw with increasing mass efficiency as the power rises. The need for very reliable control and safety systems, extensive shielding and disposal of spent or malfunctioning reactors into deep space suggests that they are likely to be applied to very high power applications where other heat sources cannot be produced with a sufficient power level. A number of closely coupled reactor and convertor types are being developed with a view to improving the overall mass efficiency. Systems which use thermionic converters in this manner have been produced in the USA and USSR. A typical reactor of this type at 10 Kw would have a mass efficiency of 60 Kg per Kw rising to 45 Kg per Kw at 25 Kw output

b) Converters

i) Thermoelectric.

A thermoelectric converter consists of a thermocouple of two dissimilar materials which generate a d.c. current when their junctions are held at different temperatures. Lead Telluride thermocouples operate with a hot temperature up to 800° K and Silicon Germanium up to 1300°K. The efficiency of a particular thermocouple is a function of the temperature difference, the electrical conductivity (q), the thermal conductivity (K) and the Seebeck coefficient (S). A figure of merit for a particular thermocouple using these parameters is given by M=S2 q/K. The conversion efficiency is of the order of 5 to 10%

ii) Thermionic Converter.

The thermionic converter is essentially a diode in which the cathode is heated by a thermal source such as a nuclear reactor. The diode operates in a space charge limited mode in which the space charge is neutralised. For a thermionic diode as used in radio the power to neutralise the charge at current densities of 1 Amp/cm-2 about 100 W. For effective thermionic power generation the space charge must be neutralised at a small energy cost.

Possible methods to accomplish this are:

- The introduction of positively charged particles into the inter-electrode space. If positive ions are introduced into the gap a plasma is formed which acts as a virtual anode very close to the cathode. For a caesium coated cathode, the ratio of ion to electron current densities needed to provide the virtual anode is 1:500.

- A second method to eliminate the space charge effect is the by the use of crossed fields. This is perhaps the most attractive of all the methods since it avoids the complications of gas filling the inter electrode space and corrosion of any other materials used by caesium. In theory, the complete elimination of the space-charge could be achieved at no electrical energy cost as shown in Figure 8.11.4. A plane cathode at zero potential and a plane parallel accelerator are separated by about 1mm for which an applied voltage of +200 V is sufficient to draw the saturated emission current. In the presence of the electric field E alone electrons emitted by the cathode would be collected at the accelerator with an accompanying loss of power from the h.t. supply.

However when a magnetic field H is applied at right angles to the electric field the system behaves like a magnetron and the emitted electrons describe cycloidal paths and are prevented from reaching the accelerator. The cycloids are given by the equations given in Figure 8.11.4.

Provided the load resistance R_L is of a suitable value, the electrons which return to be collected by an electrode at zero potential in the plane of the cathode at a distance x = 2 mE/eH2 from the origin C.

In practice the operation is much more complex and the thermal to electric efficiencies of only 10 % are currently achievable for cells based on this principle.

iii) Closed Brayton Cycle.

A closed Brayton cycle converter consists of a radial turbine, a radial compressor, an integral alternator, a recuperator and the interconnecting pipe work. The turbine may use XeHe gas as the operating fluid. It would operate at about 1400 K input temperature and have an average reject temperature of around 680° K at the waste heat radiators. The system has an efficiency of about 23% thus reducing the size of the radiator to free space. The lower waste heat temperature however may offset the efficiency gain over thermocouples.

iv) Organic Rankine Cycle

This type of converter consists of a turbine, pump, an integral alternator, a regenerative heat changer, a rotary fluid management device, a back pressure regulator, a condenser heat exchanger and a flow control valve. This converter operates at a turbine input temperature of 680° K and a mean reject temperature of 450° K with an efficiency of 18% The low waste heat temperature would require a large radiator area

v) Free Piston Stirling Engine

This system consists of a hermetically sealed unit containing a dual piston moving back and forth between two chambers. One piston is a power type, the other attached to a linear alternator which produces the electrical output. External heat exchangers

extract heat from the primary heat source and deliver residual heat to the radiators. The Stirling engine is relatively simple and two operating back to back can largely eliminate vibration or disturbances to the spacecraft. The system operates at an input and output temperatures of 1300° K and 700° K respectively and with a conversion efficiency of in excess of 20%.

vi) Alkali Metal Thermoelectric Convertor.

This also consists of a hermetically sealed vessel. It is divided into a high temperature / pressure region in contact with the prime heat source and a low temperature / pressure alumina solid electrolyte " (BASE) which has an ionic conductivity much larger than its electrical conductivity. The hot region contains liquid sodium at about 1100° K and the cold region with vapour sodium at around 800° K. A return line incorporating an electromagnetic pump recirculates the sodium working fluid. Electrical power output leads make contact with the porous positive electrode which covers the low pressure surface of the BASE and with the negative high temperature liquid sodium electrode. The over all efficiency is about 12%.

8.11.3 Energy Storage.

The requirement for energy strage arises whenever the energy source is only available intermittently if there is no local energy source on board the spacecraft. This occurs for earth satellites when they are in earth eclipse of the sun where the power system is solar, and for deep space missions when the sun is too distant to provide a sufficient energy level.

8.11.3.1 Large Thermal Mass.

This method may be employed as a means of temporary storage where the source provides thermal energy. the requirement is small and some form of thermal to electric converter is in use Whenever the local source temperature is high the method can be quite mass efficient

8.11.3.2 Batteries

The use of batteries has been the primary method of spacecraft energy storage and have been largely met by Nickel Cadmium battery cells. They do suffer from a limitation on depth and number of discharges if a long life is necessary. Nickel Hydrogen batteries have a significantly better charge and discharge characteristics and are currently in development and low level production.

Table 8.11.5 shows discharge Characteristics for geostationary and low earth orbits. For a 10 year life requirement the mass of NiH batteries is some X% lower than for NiCd.

8.11.4 Life Limiting Factors

8.11.4.1 Radiation

Power systems are prone to damage from the radiation and particle (including meteoroids) environment in space. The effect on the system is to degrade the power output performance and in the worst case to cause early failure. The degree of degradation is dependant on the orbit used and to sun spot activity during the period of operation. Solar cells which must be directly exposed to radiation from the sun and are thus suffer the greatest degradation. Electronic systems can be shielded effectively with a few millimetres of aluminium and the effect can largely be ignored. Similarly the high degree of shielding used in Radioisotopic Generators and Nuclear Reactors do not suffer from this type of damage. For simplification, the data presented lumps together degradation due to radiation and particle

a) Eccentric Orbits (e.g. TUNDRA)

Figure 8.11.3 shows the degradation in terms of the reduction in power output from the beginning of life (BOL) to the end of life EOL over a 20 year period during the interval 2000 to 2030 for InP, AlGaAs, and Si cell arrays operated in a TUNDRA orbit, a typical orbit for satellite communications at high latitudes. The pessimistic curves take into account a worst case set of solar flares during the period. The curves also illustrates the improved performance which may be obtained by using InP or GaAs cells instead of the current Si cells in arrays

b) Geostationary Orbit (GEO)

The degradation in geostationary orbit is approximately a factor of 3 less than for LEO.

8.11.4.2 Radioisotope Decay

Radioisotope based systems have lives which depend on the half life of the radio active fuels used. For example, Plutonium 238 as used in RTG's has a half life of 87.5 years. An RTG with a required output of 250 W over 5 years would need to be designed for an initial output of 290 W.

8.11.4.3 Mechanical Wear-out.

Many of the thermal to electric systems use devices with moving parts such as pumps, valves, compressors and turbines are subject to wear-out. For devices operating at temperatures in the range + 150 to − 50° C design lives of 30 years are achievable. For material operating temperatures outside this range, wear-out may become a life limiting factor.

8.11.4.4 Survivability.

a) Threat

i) Laser/thermal.

Photovoltaic systems (solar arrays) are most vulnerable to long duration laser attack, pulsed laser will cause surface damage only. The following methods provide a measure of protection

- Concentrator Arrays which narrow the field of view to the laser beam, have a higher mass which increases the thermal time constant and thermal capacity − Reactive coatings with limited reflection/transmisison characteristics. − The use of InP or GaAs cells which are capable of withstanding higher temperatures than Si cells.
- Combinations of some all of the above obviously would provide greater resistance to attack but at increasing mass and cost penalties

Nuclear Reactors and RTG's are relatively invulnerable to laser attack. The level of radiation shielding required and the consequent large mass provide a long thermal time constant. Provided the thermal conductivity of the outer casing is high only a sustained continuous beam is likely to cause damage

ii) Fragment/Particle Impact.

For all power systems it is not possible to withstand large high kinetic energy particle or fragment impact. Fragment shields are of very limited value against low energy particles

iii) EMP and RF weapons

Solar arrays incur an increase in mass to protect interconnections, cables and junctions from induced current. Nuclear systems are unaffected except in the associated control electronics which would need to incorporate standard protection devices

iv) Nuclear particle

In solar arrays the enhanced total damage fluence from nuclear electrons causes an increased 20% degradation in Si junctions

and 15° in GaAs. Oversizing of the array or increased thickness cover glasses would need to implemented. No significant effect is expected for Nuclear systems.

The total flux from neutrons are expected to have no effect on either Si or GaAs arrays. It is possible that neutron flux may have an effect on the core activity level in nuclear reactors but analysis would be required to evaluate this for specific reactors.

The effect of NPBW; Ho particles on solar arrays is to compound the radiation damage. The damage has a displacement effect which cannot be effectively annealed out. Degradation levels from this source are 20% for Si and 4% in GaAs. The effect on reactors is dependant on the specific reactor design but is not thought to be significant.

b) Counter Measures

i) Decoys

Solar arrays and the associated spacecraft has a small thermal signature and is therefore easy to mimic or to put into a dormant mode. On the other hand nuclear reactors and RTG's are large and hot and hence have a large signature. It is difficult to produce a realistic decoy and a dormant mode would be very slow to achieve if not impossible.

ii) Evasion.

A possible tactic against beam and kinetic energy weapons is evasion by spacecraft manoeuvre. For either type of power system manoeuvre would be limited by the dynamics of solar array wings or booms used to isolate reactors. Body mounted solar arrays could overcome this problem for spacecraft with low power requirements, as would body mounted reactors more generally.

8.11.5 Relative Costs

The relative costs of development and qualification, production and manufacturing facilities for low power and production is given in Table 8.11.6 against a base accounting unit X.

8.11.6 Bibliography

[8.11.1] General Purpose Heat Source-Radioisotope Thermoelectric Generator-Space Systems Division GE Valley Forge.

[8.11.2] Alternative Power Supplies For Spacecraft J.M.Harvey Marconi Space Systems. Proc. 5th European Symposium: Photovoltaic Generators in Space September/October 1986. (ESA SP-267 Nov 1986).

[8.11.3] The Status and Prospects of Indium Phosphide Solar Cells for Space Use-N.M. Pearsall, R.Hill, and A.A. Dollery. Proc. 5th European Symposium 'Photovoltaic Generator in Space'September/October 1986. (ESA SP-267 Nov 1986).

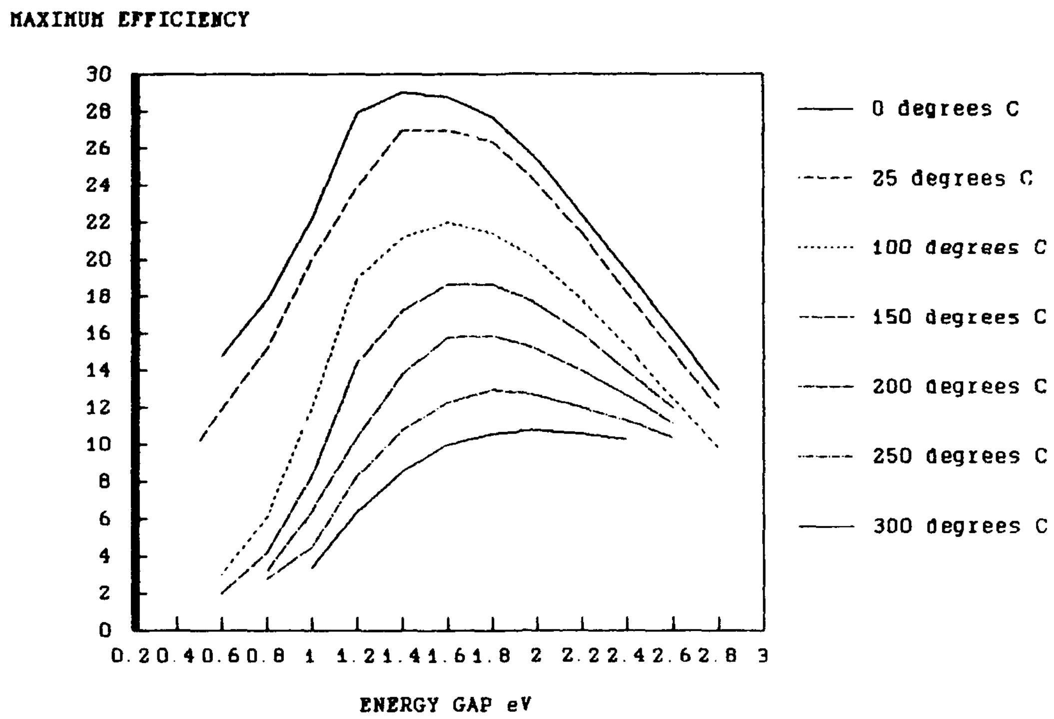


Fig.8.11.1 Maximum theoretical efficiency versus band gap eV

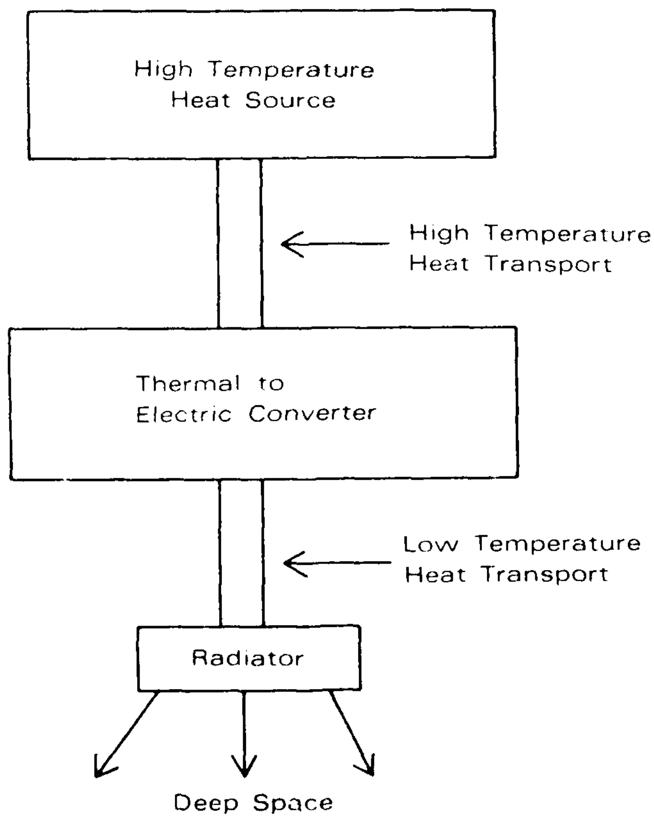


Fig.8.11.2 Thermal to electric conversion for a space mission

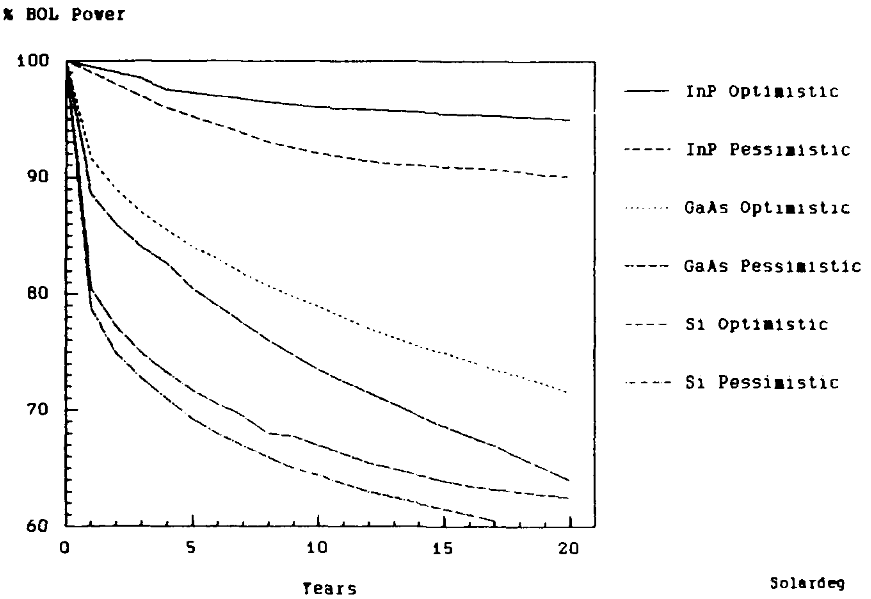
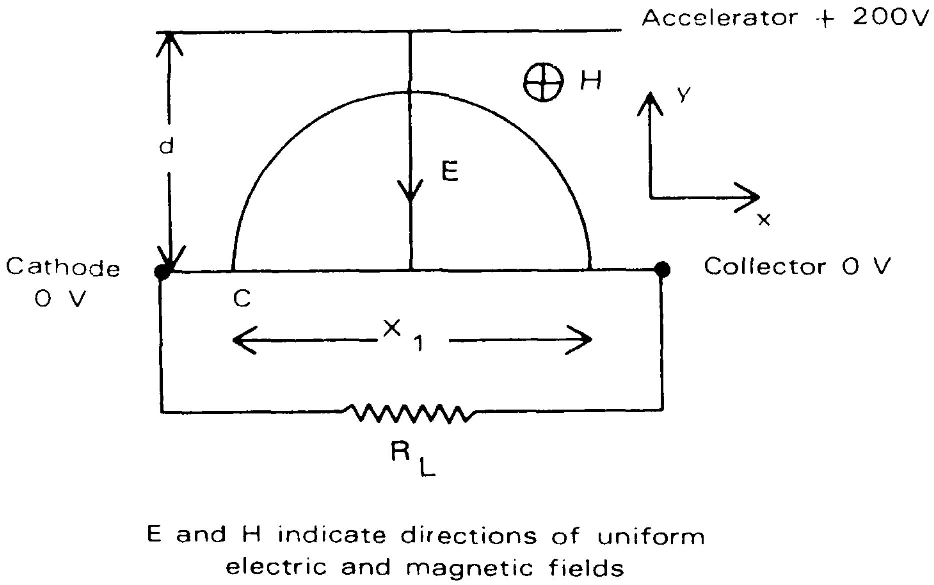


Fig.8.11.3 Solar array degradation — elliptical orbits



E and H indicate directions of uniform electric and magnetic fields

$$X = \frac{mE\,(wt - \sin wt)}{eH^2}$$

where $w = eH/m$ and $t = \text{time}$

$$Y = \frac{mE\,(1 - \cos wt)}{eH^2}$$

Fig.8.11.4 Principle of magnetic triode

Table 8.11.2
Planary array characteristics

Characteristic	Silicon	AlGaAs
Flexible array Mass Kg for 5 Kw Power Output	82	78
Rigid Array Mass Kg for 5 Kw Power Output	103	165
BOL Power Output Kw per Square Metre	0.164	0.197
EOL Power Output Kw per Square Metre	0.123	0.187
Delta Charge % over 10 years in Geostationary Orbit	- 25	- 16
BOL: Beginning of life EOL: End of Life		

Table 8.11.3
Future (10 years +) array characteristics

Characteristic	Silicon	AlGaAs
Flexible array Mass Kg for 5 Kw Power Output	40	39
Rigid Array Mass Kg for 5 Kw Power Output	103	114
BOL Power Output Kw per Square Metre	0.196	0.229
EOL Power Output Kw per Square Metre	0.142	0.196
Delta Charge % over 10 years in Geostationary Orbit	- 24	- 14
BOL: Beginning of life EOL: End of Life		

Table 8.11.4
Planar array characteristics

Characteristic	AlGaAs	InP
Array Mass Kg for 5 Kw Power Output	155	152
BOL Power Output Kw per Square Metre	0.133	0.128
EOL Power Output Kw per Square Metre	0.122	0.125
Delta Charge % over 10 years in Geostationary Orbit	- 8.3	- 2.4
BOL: Beginning of life EOL: End of Life		

Table 8.11.5
Battery characteristics 10 year life

Battery Type	Depth of Charge %	
	Geostationary	Low Earth Orbit
Nickel Cadmium	40	20
Nickel Hydrogen	70 - 80	40

Table 8.11.6
Relative costs

Systems 5 Kw	Development and Qualification	Manufacturing Facilities	Production Cost (1 off)
NUCLEAR REACTORS	5 0 % to 100 %	10 %	20 %
PHOTOVOLTAIC TODAY			
(Rigid and Flexible)			
Silicon	0 0 %	0 8 %	1 0 %
Gallium Arsenide	0 2 % to 0 4 %	0 5 %	1.5 % to 1.6 %
PHOTOVOLTAIC · 10 Yrs			
(Rigid and Flexible)			
Silicon	0 1 %	0.05 %	1 0 %
Gallium Arsenide	0 2 %	0 05 %	1 2 %
(Concentrators)			
Gallium Arsenide	2 0 %	2 0 %	2.0 %
Indium Phosphide	10 0 %	2 0 %	2 1 %

8.12 SPACECRAFT PROPULSION SYSTEMS

8.12.1 Introduction

Propulsion systems are used in the process of changing or maintaining orbit, changing or maintaining spacecraft attitude, and to change station. Change of orbit usually occurs as a phase in launch and injection to the final operating orbit. A typical case is the circularisation of the initial final elliptical launch orbit into a geostationary orbit. Maintaining orbit or spacecraft attitude is required so that an accurate place or position is retained in the presence of perturbances caused by e.g. the non-spherical Earth, the Moon or the seasons. Change of station is usually only required if the area of operations covered by the spacecraft is radically altered. This may occur when a companion spacecraft fails and the spacecraft is required under an inter-operability agreement to provide a limited mutual service.

The systems fall into the following categories;

a) Solid Propellant.
b) Cold Gas.
c) Mono-propellant.
d) Bi-propellant.
e) Electronic.

The system which is chosen will depend on the which mission requirements are best met from the list below

a) Long Life.
b) Re-usability.
c) In-orbit operations such as servicing, docking and rendezvous.
d) Man-rating.
e) High energy.
f) Inter-planetary mission.

8.12.2 Solid Propellant

This system is only used for changing orbit after the initial injection phase of the launch. It is only economic where a one slot large manoeuvre is needed. It has to be supplemented by another system where fine adjustment of orbit is needed. It has largely been supplanted by bi-propellant systems which have much greater versatility.

8.12.3 Cold Gas Systems

This type of system is ideal for fine attitude control, docking and rendezvous. The propellant is either compressed nitrogen or liquid propane, both of which are non-reactive e.g. with optical devices. The system has a long life with adequate propellant storage and is man-rated.

This system is currently adequate for purpose and is not likely to changed if lower weight materials become available for the storage pressure vessels, thrust orifices and valves.

8.12.4 Mono-Propollent

Mono-propellant systems are simple, well developed, reliable and cheap. The fuel used is hydrazine which is disassociated by a catalyst, and evolves heat. The system can be used many times and is therefore suitable for both attitude and orbit control and for station change. It is usually operated in a pulse mode, the number of pulses being proportional to the total impulse required. The life time is limited by thrust chamber erosion. Improvements in this respect may arise from developments in materials.

Hydrazine systems are unsuitable for manned operations in orbit due to the extremely high toxicity. Hydrazine is also highly reactive and care is required in spacecraft design to avoid materials problems. Thrust levels in the range 0.5N to 10N are currently available.

8.12.5 Bi-Propellant Systems

This system uses two fuel components - mono-methly hydrazine and nitrogen tetroxide. Thrust levels of between 5N and 20N are possible, and the system is 40% more mass efficient than an equivalent mono-propellant system. The system components taken alone are more expensive than mono-propellant systems but the lower mass reduces the overall costs by a larger amount. The combustion chamber of bi-propellant systems run hotter than the mono system and erosion is faster as a consequence. New materials are being investigated and are expected to improve the life characteristics.

The system is similar to the mono-propellant one in that it can be used many times. Its mass efficiency allows it to undertake orbit change functions as well as attitude and orbit control functions, and is replacing the solid propellant Apogee Kick Motor (AKM) previously used for this operation.

8.12.6 Dual Mode Systems

This system uses a bi-propellant system in two modes.

1) Bi-propellant
2) Mono-propellant

For both cases the hydrazine used is the same i.e. standard hydrazine. Thrust levels are currently limited to not less than 1ON but 5N thrusters are under development.

8.12.7 Electronic or Ion Thrusters.

These systems will provide thrust levels in the range 10 -50 mN and are suitable for station keeping, drag compensation in low earth orbit, and very suitable for long duration interplanetary mission orbit correction and manoeuvre. They are particularly useful for maintaining very fine pointing accuracies and for de-orbiting spent spacecraft where a small thrust level can be maintained for a very long time. Development is at a relatively early stage but their use is currently planned for an ESA Technology Mission (SAT 2) in an operational role.

8.13 SPACE TRANSPORTATION

8.13.1 Introduction

In the period 2000 to 2030 Space transport systems are expected to develop towards a range of expendable vehicles which are an economic match to spacecraft mass volumes and orbital requirements.

Currently the number of communication satellites which operate in geo-stationary orbit provide sufficient traffic which make large launchers carrying multiple satellite payloads the most economic. At the same time the number of satellites which are required in non-geostationary orbits of the same type is very small and the possibility of multipel payloads thus very infrequent. Smaller capacity launch vehicles are thus a better match to satellite mass and volume and are more economic.

As the overcrowding of the geo-stationary orbit increases more communication satellites will adopt Molniya or other elliptical orbits, and the traffic around 2000 in these orbits may well approach the current geo-stationary levels and large launchers again be the most economic.

For military communication satellites however, continuity of service in times of stress or war is vital and it will be necessary to replace any defective satellite or one which has been lost by enemy attack as quickly as possible. It is then necessary to have available both a suitable launcher and launch site. It will not then be sensible to wait for a suitable partner satellite to become available for a multiple payload launch.

At present the number of launch sites are few, and are fixed geographically. These are also vulnerable and are likely to prime targets for an enemy attack.

8.13.2 Requirements For Launchers and Launch Facilities

During peacetime conditions the availability of launchers and facilities generally is not a problem. The only restriction being the use of facilities with adequate security arrangements. Delays in the replacement of failed satellites can be accepted since inter-operability agreements can together with the networks provide adequate interim communications capacity.

During periods of crisis or war, access to the full range of these peacetime facilities may be denied for political reasons, and delays in th replacement of satellites and launch facilities which have been destroyed by enemy action can no longer be tolerated. There then maybe a case for a stock pile of dedicated launch vehicles and the provision of a series of dispersed and hardened or mobile launch facilities. The system space and ground architectures will determine the extent of these provisions.

8.13.3 Launch Vehicles

A full listing of current international launch vehicles is given in Appendix 8.13 A. Not all of these would be accessible for NATO purposes in times of crisis or conflict and many will be superseded by 2000-2030.

Figure 8.13.1 shows a selection of current launchers and future vehicles expected to be available in the period 2000 to 2030. This list is likely to change during the interval to the year 2000, in particular with respect to smaller capacity launchers.

The Pegasus booster is a particular example of a small launcher (Appendix 8.13B). This vehicle is capable of putting a 270 Kg payload into a 400 Km polar orbit. It is carried by a converted transport aircraft (eg Tristar) and launched from it. It is claimed to cost one half of a launch from the ground since it makes use of aerodynamic lift given by the aircraft wings and by the use of air as the fuel oxidant during the first stage of flight. Launch from the air gives great mobility and short duration launch preparations of the order of days rather than the weeks of conventional ground based launches. Its main limitations are payload mass and limited orbit height. Further developments in this area can be expected from carriers such as Hotol, Sanger and Hermes for he first stage which could extend both the payload mass and orbit altitude capability.

8.13.4 Vehicles For Repair Or Recovery

The use of space transport in-orbit repair or recovery for repair on the ground at first sight appears attractive. Using the STS (Shuttle) this has certainly been shown to be feasible provided the spacecraft to be reparied or recovered can be brought down to an accessible orbit. It has however not been shown to be truly economic. The problem becomes more severe when the spacecraft is in geostationary or highly elliptical orbit as the cost of the fuel used to either move the satellite to the recovery vehicle orbit or move an inter-orbit transfer vehicle to the satellite orbit and back becomes at least as great as the cost of a new satellite and launch. Repair and recovery in which man plays an intimate in orbit part is even more expensive. The extensive use of man in orbits which spend a significant time in the Van Allen belts(e.g. Molniya) is impracticable due to the hazards of exposure to the high radiation levels.

In the time frame 2000 to 2030 mans involvement is likely to be confined to the in-orbit construction and servicing of Space Stations and other large structures such as real aperture Radars, at altitudes less than 500 Km.

8.13.5 Launch Sites

The Launch sites which are likely to be available for the launch of NATO satellites are Cape Canaveral, USAF Vandenberg and the ESA site at Kirou. These sites are obvious prime targets for ICBM's or submarine launched Cruise missiles. Consideration should be given to the provision of less vulnerable alternatives such as mobile launch platforms or a multiplicity of hardened launch pads such as the Minuteman silos. Mobile launch pads for small satellites of masses, fixed hardened silos would be more suitable. The necessary stockpile of launch vehicles could be made less vulnerable still, by dispersed storage elsewhere. For very small payloads the air launched Pegasus launch vehicle would provide a range of mobility only limited by the availability of suitable airfields for the aircraft first stage.

8.13.6 Spacecraft Control Stations

The control of satellite launch currently requires a considerable degree of down range tracking. Several geographically widely separated ground stations are used during this phase. When in the required orbit, control operations can be either autonomously executed by the satellite or via the normal communications up and down links using a large range of use would reduce the reliance on skilled technical manpower who would themselves be prone to attack. The autonomy aspect is dealt with further in section 8.10 Robotics and Control.

APPENDIX 8-13A

US LAUNCH VEHICLES

Vehicle Contractor/ Vehicle Name	User Agency	PROPULSION Stage No	Engines	Stage Contractor	Stage or Motor Designation	Propellants oxidizer/fuel	Thrust (lb)	DIMENSIONS & WEIGHT Max Dia (ft)*	Length (ft)**	Launch Weight (lb)	PERFORMANCE Payload (lb) Orbital	Escape
BASIC VEHICLES												
Martin Marietta												
Titan 34D Transage	USAF	0	2x1120 in UA 1205 (strap-on)	UTC	-	Solid	246.288.000*	10.2	91.4	1.514.600	4.700*	
		1	2xAerojet LR-87-AJ-11	Martin Marietta	-	N_2O_4/N_2H_4-UDMH	529.000	10.0	78.3			
		2	1xAerojet LR-91-AJ-11	Martin Marietta	-	N_2O_4/N_2H_4-UDMH	101.000	10.0	37.0			
		3	2xAerojet AJ10-138	Martin Marietta	Transtage	N_2O_4/N_2H_4-UDMH	16.000	10.0	14.7			
Titan 34D No Upper Stage	USAF	0	2x120 in UA 1205 (strap-on)	UTC	-	Solid	246.788.000*	10.2	91.4	1.492.200	27.600	
		1	2xAerojet LR-87-AJ-11	Martin Marietta	-	N_2P_4/N_2H_4-UDMH	529.000	10.0	76.6			
		2	1xAerojet LR-91-AJ-11	Martin Marietta	-	N_2P_4/N_2H_4-UDMH	101.000	10.0	31.3			
Titan 2SLV No Upper Stage	USAF	1	2xAerojet LR-87-AJ-5	Martin Marietta	-	N_2O_4/N_2H_4-UDMH	430.000.**	10.0	71.2	340.000	4.200	
			1xAerojet LR-91-AJ-5	Martin Marietta	-	N_2O_4/N_2H_4-UDMH	100.000(Vac)	10.0	23.4			
Titan 3	Commercial	0	2x120 in UA 1205 (strap-on)	UTC	-	Solid	246.298.000*	10.2	91.4	1.492.200	27.000	
		1	2xAerojet LR-87-AJ-11	Martin Marietta	-	N_2O_4/N_2H_4-UDMH	529.000	10.0	71.6			
		2	1xAerojet LR-91-AJ-11	Martin Marietta	-	N_2O_4/N_2H_4-UDMH	101.000	10.0	31.3			
Titan 4 Centaur G Prime	USAF	0	2x120 in UA 1207 (strap-on)	UTC	-	Solid	319.400.000*	10.2	112.9	1.910.409	10.000	
		1	2xAerojet LR-87-AJ-11	Martin Marietta	-	N_2O_4/N_2H_4-UDMH	546.000	10.6	86.5			
		2	1xAerojet LR-91-AJ-11	Martin Marietta	-	N_2O_4/N_2H_4-UDMH	104 000	10.0	37.6			
		3	2x P & W RL 10A-3A3	GD Space Systems	-	LOX/LH_2	33.000	14.2	29.3			
Titan 4 105	USAF											
GD/Space Systems												
Atlas G Centaur	NASA	1	2xRockeldyne YLR-89-NA7	GD/Convair	MA-5	LOX/RP-1	377.500	10.0	140 5**	360 000	5.200**	3 501
D-1A/Atlas H		1	1xRockeldyne YLR-105-NA7	GD/Convair	-	LOX/RP-1	60.000	-	104.7	293.000	3.000**	
McDonnell Douglas												
Delta 3920 PAM-D21	NASA	1	1xRocketdyne RS-27	McD/Douglas	ELT Thor	LOX/RP1	231 600	8	74.8	423 000	2 830	2 000
		1	9xThiokol TX526-2	McD/Douglas	Castor 4	Solid	801.400*	3.3	36 6	429.000	-	-
		2	1xTRW TR201/1xAerojet AJ10-118K	McD/Douglas	Delta	N_2O_4/N_2H_4-UDMH	9.850/10.000	8	19.3			
		3	1xThiokol Star 48	McD/Douglas	PAM-D	Solid	15.000	4	6.7			
Delta 6920 (has 1st two stages only)	USAF	1	1xRocketdyne RS-27	McD/Douglas	Extra ELT Thor	LOX/RP1	241.000	0	86 8	483 700	3.190**	2 270
Delta 7925(3 stages)		1	9xThiokol TX-780	McD/Douglas	Castor 4A	Solid	929.400	3.3	36.6			
		2	1xAerojet AJ10-118K	McD/Douglas		N_2O_4/N_2H_4-UDMH	10.000	8	19.6			
		3	1xThiokol Star 488	McD/Douglas	PAM-D	Solid	15.000	4	6 7			
Delta 7920 (has 1st two stages only)	USAF	1	1xRocketdyne RS-27	McD/Douglas	Extra ELT Thor	LOX/RP1	241.000	8	87.4	511.000	4.000**	2 814
Delta 7925(3 stages)		1	9xHercules GEM	McD/Douglas	Gr-Ep Motor(GEM)	Solid	985 500	3.3	42.6			
		2	1xAerojet AJ10-118K	McD/Douglas	N_2O_4/N_2H_4-UDMH		10 000	8	19 6			
		3	1xThiokol Star 488	McD/Douglas	PAM-D	Solid	15.000	4	6 7			
Vought												
Scout SLV-1A	NASA	1	1xUTC Algol 3	LTV	Algol 3A	Solid	107 000	3.7	75.1	47.200	400**	75
	USAF	2	1xThiokol Castor 2	LTV	Castor 2A	Solid	61 000	-	-			
		3	1xThiokol Star 31	LTV	Antares 3A	Solid	21 000	-	-			
		4	1xThiokol Star 20	LTV	Altair 3	Solid	5.700	-	-			
UPPER STAGES												
GD/Space Systems												
Centaur D-1A/D-1T**	NASA	Varies	2xP&W RL 10A-3-3A	GD/Convair	Centaur	LOX/LH_2	33.000	10 6	30.0	35 000	5 200** 17.500**	1 500 13.000
Martin Marietta												
Transtage	USAF	Varies	2xAerojet AJ10-138	Martin Marietta	Transtage	N_2O_4/N_2H_4-UDMH	16 000	10	15 8	27 000	4 200*	4 000
Fairchild/Space												
Stage Vehicle Sys	USAF	2	2xThiokol Star 37	Fairchild/Space	SGS Blk I	Solid	15 500	4 6	10 3	3 500	-	
Orbit Insertion Sys	USAF	1	1xThiokol Star 27	Fairchild/Space	OIS	Solid	6.000	4.6	5.9	1.200	-	-
McDonnell Douglas												
Stage Vehicle Sys (SGS-III)	USAF	1-2	2xThiokol Star-48	McD/Douglas	SGS-2	Solid	16.000	4.5	13 0	11.700	1 000**	-
STS/PAM-A	NASA	1	1xThiokol (MM3)	McD/Douglas	PAM-A	Solid	35.200	5 0	7.5	12.700	4.400**	2 530
STS/PAM-D	Varies	Varies	1xThiokol Star-48	McD/Douglas	PAM-D	Solid	15.000	4.3	6.3	7.000	2 750**	1 630
STS/PAM-DII	Varies	1	1xThiokol Star 630	McD/Douglas	PAM-DII	Solid	17.600	5 3	6 5	12.000	4.000	2.300
Boeing												
IUS	USAF NASA	1-2	SRM-1 SRM-2	Boeing	SRM-1 SRM-2	Solid Solid	44 100 16 000	9.5	16.0	32.300	5 000-6 000	11 523** 3 30**
Orbital Sciences												
Transfer Orbit Stage	Varies	Varies	SRM-1	Orbital Sciences	TCS	Solid	44.100	9 6	10.7	24.000	13.400**	7.930
Apogee and Manoevering Stage	Varies	Varies	Rocketdyne RS-51	Orbital Sciences	AMS	N_2O_4/MMH	2.650	12 0	5 4	11.200	5 600**	2.890
TOS/AMS	Varies	Varies	SRM-1 Rocketdyne RS-51	Orbital Sciences	TOS/AMS	Solid N_2O_4/MMH	44.100 2.650	12.0	15.7	35.300	6.500	10.100

See International Launch Vehicles for abbreviations and notes

APPENDIX 9.13 A

INTERNATIONAL LAUNCH VEHICLES

Country User Agency/ Vehicle Name	Vehicle Contractor	Propulsion Stage No	Engines	Stage Contractor	Stage or Motor Designation	Propellants (oxidizer/fuel)	Thrust (lb)	Dimensions & Weight: Max Dia (ft)*	Length (ft)**	Launch Weight (lb)	Performance Payload (lb): Orbital	Escape
PEOPLE'S REPUBLIC OF CHINA												
FB-1 (CSL-2)**	-	1	4 x-	-	-	N_2O_4/UDMH	617 300	10.9	68.3	420.000	4.4x10*	-
		2	1 x-	-	-	N_2O_4/UDMH	154.300	10.9	39.2			
CZ-3	-	3	-	-	-	inc LOX/LH_2 stage(s)	-	-	-	-	-	Heavy Payloads
FRANCE												
ESA/arianespace												
Ariane 2	CNES/Arianespace	1	4 x Viking 5 liquid	Aerospatiale/SEP	L-140	N_2O_4/UH25	601.000	12.5	59.8	490.000 (total)	4.795 (geostationary transfer	3 200**
		2	1 x Viking 4 liquid	ERNOSEP	L-33	N_2O_4/UH25	177.600	8.5	37.6			
		3	1 x HM-7B liquid	Aerospatiale/SEP	H-10	LOX/LH_2	14.000	8.5	34.2			
Ariane 3	CNES/Arianespace	1	4 x Viking 5 liquid	Aerospatiale/SEP	L-140	N_2O_4/UH25	601.000	12.5	59.8	530 000 (total)	5.690 (geostationary transfer	3.790**
		1	2 x P7.3 solid	BPD	PAP	Solid	250.000	3.5	26.2			
		2	1 x Viking 4 liquid	ERNOSEP	L-33	N_2O_4/UH25	117.600	8.5	37.6			
		3	1 x HM-7B liquid	Aerospatiale/SEP	H-10	LOX/LH_2	14.000	8.5	34.2			
Ariane 4?*	CNES/Arianespace Arianespace	1	4 x Viking 5 liquid	Aerospatiale/SEP	L-220	N_2O_4/UH25	601.000	12.5	82.5	523.000 to 1.833.000	4.190 to 9.260 (geostationary transfer 7° incl)	-
		1*	2-4 x Viking 6 liquid	ERNOSEP	L-36	N_2O_4/UH25	152.000	7.1	62.3			
		1*	2-4 x P9.5 solid	BPD	P9.5	Solid	146.000	3.5	39.2			
		2	1 x Viking 4 liquid	ERNOSEP	L-34	N_2O_4/UH25	177.000	8.5	37.6			
		3	1 x HM-7B liquid	Aerospatiale/SEP	H-10	LOX/LH_2	14.000	8.5	34.2			
INDIA Indian Space Research Organisation (ISRO)												
SLV-3	VSSC	1	1 x solid (S-1)	VSSC	-	Solid	95 000	-	74.5	37.500	80	-
		2	1 x solid (S-2)	VSSC	-	Solid	-	-				
		3	1 x solid (S-3)	VSSC	-	Solid	-	-				
		4	1 x solid (S-4)	VSSC	-	Solid	-	-				
JAPAN												
National Space Development Agency (NASDA)												
N-2**	MHI	1	1 x Rocketdyne MB-3	MHI Stel	DSV-3P-1	LOX/RP-1	172.000	8.0	74.5	297.600	4.400	770
		1	9 x Thiokol TX354-5	NM	Castor 2	Solid	52.000 ea	2.6	23.8			
		2	1 x Aerojet AJ10-118F	MHI Stel		N_2O_4/Aerozine 50	10.000	8.0	19.0			
		3	1 x Thiokol TE364-4	MHI Stel		Solid	15.000	8.0	5.7			
H-IA	MHI	1	1 x Rocketdyne MB-3	MHI Stel	DSV-3P-1	LOX/RP-1	172.000	8.0	74.5	306.460	7.100	-
		1	9 x Thiokol TX354-5	NM	Castor 2	Solid	52.000 ea	2.6	23.8			
		2	1 x LE-5	MHI Stel	-	LOX/LH_2	22.000	8.0	26.2			
		3	1 x UM-129A	NM	-	Solid	17.600	8.0	8.5			
H-2	MHI	1	1 x -	MHI Stel	LE-X	LOX/LH_2	209.000	13.1				
		1	2 x -	NM	-	Solid	23.100	5.9	150.9	528.000	4.410	-
		2	1 x -	MHI	LE-5	LOX-LH_2	286.000	13.1				
ISAS												
Mu-3S-II	NM	1	1 x -	NM	M-13	Solid	283.800	4.6	47.7			
		1	2 x -	NM	SB-735	Solid	73.700	2.4	29.9	136.400	1.700	304
		2	1 x	NM	M-23	Solid	117.500	4.6	20.9			
		3	1 x -	NM	M-3B	Solid	29.700	5.4	22.5			
USSR												
Soyuz** (SL-4)	-	1	16 x RD-107	-	RD-107	LOX/Kerosene	900.000	33.8	62.8	720.000 (total)	16 500	-
		1	4 x RD-108	-	RD-108	LOX/Kerosene	225.000	9.8	91.8			
		2	4 x RD-108	-	RD-108	LOX/Kerosene	225 000	9.8	32.8			
Proton** (SL-9)	-	1	6 x liquid propellant	-	-	LOX/UDMH	-	-	-	-	43.000	
		1	liquid propellant	-	-	LOX/UDMH	-	-	-			
		2	liquid propellant	-	-	LOX/UDMH	-	-	-			
SL-13**	-	0	6 x RD-253	-	RD-253	N_2O_4/N_2H_4-UDMH	-	-	-	-	50.000	
		1	-	-	-	LOX/UDMH	-	-	-			
		2	-	-	-	N_2O_4/N_2H_4-UDMH	-	-	-			
		3										
Energia		stages	4 liquid strap-ons	-	-	-	-	-	12.0	400 000	200.000 to 5.000.000	

*Excluding strap-ons **Excluding payload

Abbreviations

CNES - French National Center for Space Studies	ISAS - Inst of Space & Aero Sciences-U of Tokyo	MPC - Mitsubishi Precision Co
ESA - European Space Agency	ISRO - Indian Space Research Organisation	NEC - Nippon Electric Co
CD - General Dynamics	LOX - Liquid oxygen	NOF - Nippon Oil and Fat Co
IRFN - Inhibited red fuming nitric acid	McD/Douglas - McDonnell Douglas	NASDA - Japanese National Space Development Agency
RP1 - Hydrocarbon rocket prop	MCI - Matsushita Communication Industries	NM - Nissan Motors
SEP - Societe Europeenne de Propulsion	MHI - Mitsubishi Heavy Industries	P&W - Pratt & Whitney Aircraft
	SLV - Standard Launch Vehicle	UTC - United Technologies Corp Chemical Systems
	UDMH -	VSSC - Vikram S. Space Center

1 100-naut mi(115-stat-mi)circular orbit
2 Titan 34B D Titan 2 SLV from Western Test Range 34D Transtage/4 from Eastern Test Range
3 Total trust of strap-on rockets is shown.All Deltas ignite six at lift-off three later
4 Total thrust in lb sec generated during solid burn
5 Synchronous-Equatorial orbit
6 Strap-on liquid or solid boosters Ariane 4 series
8 Fairchild stage vehicle system with Navstar payload for placement into 100)10.900 naut-mi transfer orbit
10 300 naut-mi(345-stat-mi)orbit from due-east launch
11 Multi-burn capability Bask vehicles include Titan 3B Atlas SLV-3A and Thor payload shown are with Atlas SLV-3A booster
12 Centaur D-1A stager are for Atlas SLV-3D vehicles Centaur D-1T is a modified Centaur used in connection with Titan 3E Both have multi-burn capabilities the D-1 T has demonstrated a zero-g to restart capability as well
13 Developed from SS-6 ICBM "SL" is "US" designation for Soviet spacecraft launchers
14 Basic civilian launch vehicle
15 Geo-synch transfer orbit Atlas E Booster
16 Salyut launcher
17 Geo synch transfer orbit with Titan T-34D
18 Length including none fairing
19 450 nm circular sunt synch orbit with AKM
20 Improved N-1 with Delta Guildance system planned for 1980-81 launch Improved N-2 with a cryogenic stage also under study
21 PAM-D is payload Assist Module-Delta class, a McDonnell Douglas convercial development capable of operating either as a third stage of Delta or spin-stabilized upper stage from shuttle
23 Two stage escape
24 Two stage planetary
25 Synch transfer orbit
26 US designation
27 200
29 Transfer orbit for a 12-h final orbit period
30 Vinf = ONCH'S
31 Limited by lift capability
32 Geostationary orbit
33 450 nm circular sun synch. orbit with AKM
34 For Development Launch
35 For operational Launches
36 Ariane 4 is a family of Launchers designated AR4
37 Upper stage option
38 Sea level

APPENDIX 8. 13B

AIR-LAUNCHED SPACE BOOSTER (PEGASUS)

1. Pegasus is an air-launched space booster (Fig. 1) with the following capabilities:

   - Weight : 41,000 Lb
   - Length : 50 Ft
   - Diameter : 50 In
   - Wing Span : 22 Ft
   - Payload : 600 Lb to 250 nmi Polar

2. Pegasus is a commercially developed launch vehicle and will initially use a converted B-52 aircraft and later converted commercial transport aircraft will be employed.

3. The following advantages are claimed for Pegasus:

   a) AFFORDABLE
      i) One-half cost of ground launch
      ii) Schedule assurance
      iii) Reflight guarantee

   b) DEPENDABLE
      i) Dedicated launch
      ii) Weather insensivity
      iii) Off-the-shelf launcher

   c) VERSATILE
      i) Launch origin
      ii) Orbit selection

4. The Pegasus is said to have the payload delivery capability of an identical ground-launched vehicle for the following reasons:

   - Energy imparted by the carrier aircraft
   - Reduced aerodynamic drag losses
   - Improved propulsion efficiency
   - Reduced gravity losses
   - Reduced thrust direction losses

5. The operational benefits of air-launched boosters are:

   a) Omni-inclination launch
   (Mission flexibility, eliminates doglegs, performance improvement)
   b) Reduced constraints on launch
   (Fly above or around weather, no conventional launch pad costs or bottlenecks, reduced range safety concerns)
   c) Simplified Ground Operations and facilities
   (Airport-based integration and flight operations, horizontal buildup, simple vehicle handling equipment)
   d) High altitude hypervelocity missions
   (propulsion, materials and flight dynamics experiments for hypersonic vehicles)

6. Pegasus appears to be the only commercial mobile space launch vehicle with the attributes given below:

   a) Autonomous mission operations; aircraft provides release at launch point and approximate heading.
   b) Payload injection into final orbit within line-of-sight of aircraft 12 minutes after release.
   c) Aircraft can provide autonomous range support activities (tracking, telemetry and flight safety)
   d) Pegasus can be assembled and tested within two weeks because:

      i) Horizontal integration using V-rail and shipping integrating dollies eliminates need for overhead cranes.
      ii) Factory pre assembly and pretesting of major subsystems reduces launch site operations.
      iii) Use of solid propellant (class 1.3) and cold gas propulsion systems reduces handling complexity and enhances safety.

7. Pegasus is mated on launch day because:

   a) Mating requires approximately 4 hours.
   b) Assembly trailer provides vehicle for aircraft mating.

8. It is stated by the manufacturer that Pegasus has been static tested and has the following launch schedule:

   a) First program year: every 3 months
      - 1989-August, October
      - 1990-January, April.
   b) Second program year: every 2 months
      - 1990- June, August, October, December
      - 1991- February, April
   c) Third program year: every month
      - 1991- May, June, July, August, etc.

Table 8.13.1
Launch vehicle availability

LAUNCH VEHICLE	LIFT CAPABILITY (Kg)				COST/Kg		YEAR
	LEO Polar		GTO		LEO	GTO	1990 2000 2010 2020 2030
	MIN	MAX	MIN	MAX			
CURRENT VEHICLES							
DELTA 6925	2000	2600	1270	1500	10000	25000	
ATLAS CENTAUR	3200	5000	2000	2360	10000	25000	
TITAN III	10000	14000	1270	2600	10000	25000	
ARIANE 3	3000	4500	1200	2580	10000	-	
ARIANE 4	3800	7500	1000	4200	10000	-	
SHUTLE	-	30000	-	-	13000	25000	
FUTURE VEHICLES							
HERMES	-	4500	-	-	-	-	
HOTOL/SANGER	-	4000	-	-	4000	8000	
ARIANE 5	-	10000	2200	5200	12000	20000	
MLV	-	6000	-	-	5000	-	
HLV	-	45000	-	-	5000	-	
ALS	-	68000	-	-	5000	-	
DELTA 7925	2500	3200	1500	1800	8000	-	
PEGASUS	-	270	-	-	2500	-	

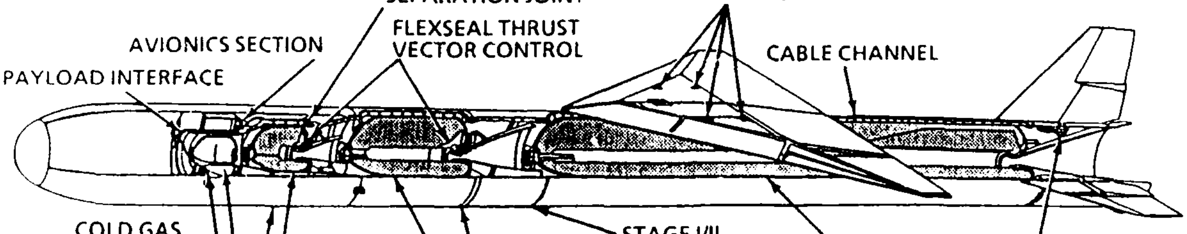


Fig.1 Pegasus launch vehicle

8.14 EARTH STATION TECHNOLOGY

8.14.1 Technological Aspects

The earth stations will, to a large degree, depend on technology in use or under development today. The trends will continue towards more easily transportable or even man-packed terminals for some of NATO requirements. This will depend on more capable satellites with EIRP similar to civilian satellites for VSAT networks. To combine this trend with useful AJ capability will be a challenge both for satellite technology and for ground terminal technology.

Efficient solid-state transmitters at EHF will be an important technology for ground terminals. In general the earth terminals will benefit from progress in digital and analog (MMIC) solid state technology.

The uplink frequency hopping signal used to access the processing transponder (see. Sec. 8.2) can be generated by one or possibly two synthesisers per access. Small terminals will have only one access and compact synthesisers may be the preferred technology. To maintain earth terminal simplicity, it seems necessary that the hopping generator is free running and that the fine hopping synchronisation is made on board the spacecraft separately for each access. The means for doing this is indicated in Section 8.2.

Terminals with a large number of uplink carriers can generate all carriers including their frequecny hopping by using the Inverse of the Chirp Fourier Transform (ICFT) described in Section 8.2. The principles of the ICFT are given in App. 8.2A. The ICFT will, by using two physical filters, allow all carriers to hop over a wide bant (suggested to be 200 MHz in sec.8.2). To allow hopping over the full 2 GHz uplink band several ICFT's can be frequency multiplexed or one ICFT may be augmented by mixing the ICFT output signal with a coarsly hopping synthesiser signal.

Active phased arrays are strong candidades as antennas both for small and large ground terminals. Phased arrays will offer advantages over fixed reflector antennas in beam pointing agility,trasportability and in easier camouflage. Continued development is expected to better electrical properties and lower costs by addressing the technical issues discussed in Appendix 8.3 A. Current cost levels for terminals indicated in Fig.8.14.1 will not necessarily be exceeded for the more capable future terminals.

If the receiving space segment consists of a cloud of satellites extending over an angle exceeding the beamwindth of the jammer,an AJ advantage may be realized by spatial hopping of the uplink beam. To exploit this possibility,phased arrays are essential for the uplink transmitter to provide the beam agility required.

Active phased arrays can be small and compact.A 1600-element active array would give 200 mW at 44 GHz is to be expected within 10 to 15 years (this capability is currently available at 20 GHz).Such a 1600-element active transmit array antenna would provide 60 dBW EIRP. The total size of the 44 GHz receive band will have a size of 30cmx30cm.An optimistic cost for an element of such an array antenna would be about 100 US$.

An alternative to active array antennas would be to use a mechanically steered dish antenna.A suitable dish in a 50 cm radome driven by a 100 W TWTA at 40 GHz would provide 60 dBW EIRP. The cost would be lower than for the phased array.

MMIC (GaAs) are important building blocks to realize light-weight and affordable electrical conductors to distribute signals to/from the radiating elements.

8.14.2 Mobile Earth Terminals

8.14.2.1 System Aspects

Military networks with mobile earth stations are primarily designed either to collect or distribute data to or from a number of remote terminals from a central hub station.Alternatively,an interactive two-way network can be established,usually on a TDMA basis.

Mobile terminals for military satellite communication are characterized by the following:

- Large number of users
- Geographically dispersed
- Transmission of short messages
- Low duty cycle per user
- Low antenna gain
- Small and easy to transport
- Rugged construction
- Low cost.

The so-called VSATs (Very Small Aperture Terminals), developed and used so far mainly for civil communications task or perform the essential mililtary communications task of transmitting short messages,e.g. data or voice packets. These terminals are of a simple and rugged construction,further on they are small and therefore easy to transport;for the abovementioned reasons the VSATs are extremely appropriate for changing battlefield conditions, they also enable a fast deployment of forces

Future satellites are increasingly incorporating multiple beams for direct and effective energy distribution with high EIRP and G/T which results in smaller,more economical earth terminals.For the provision of direct VSAT-to-VSAT service,these satellites could also be equipped with onboard signal processing,an approach that makes the most efficien use of the transponder's available power and bandwidth for VSAT use.

For the existing Defense Commucations Satellite System (DSCS),considerations about VSATs as required by the Air Force have been performed.This new demand for easily transportable and deployable satellite terminals cannot be met by the SHF terminals used so far.A transmission analysis resulted in VSAT designs with reflector antennas of about 2 feet diameter,typically using SHF power amplifiers with outputs from 1 to 4 watts for low bit rate users(about 2400 bits/s)

8.14.2.2 Antennas

As an alternative to reflector type antennas used so far for mobile terminals, microstrip patch array antennas are wellsuited for this application.The main advantages of these microstrip antennas are;

- Quasi two-dimensional geometry with a low profile
- Small dimensions and less bulky than a comparable paraboloid
- Physically robust.

A major problem for satellite-based land mobile vehicle

communications is the continously changing reception situation scenario with related large field-strength variations due to multipath effect (signal reflection, refraction, and scattering), signal fading and shadowing by buildings or vegetation. Further on, one has to deal with unwanted signals, e.g. those of other satellites or terrestrial sources, in the case of war time possibly emited by hostile jammers.

Civil land mobile communication terminals are not without reserve suited for military purposes: the conventional antenna system with a quasi-omnidirectional pattern in most cases is too susceptible to interference caused by natural or artificial sources.

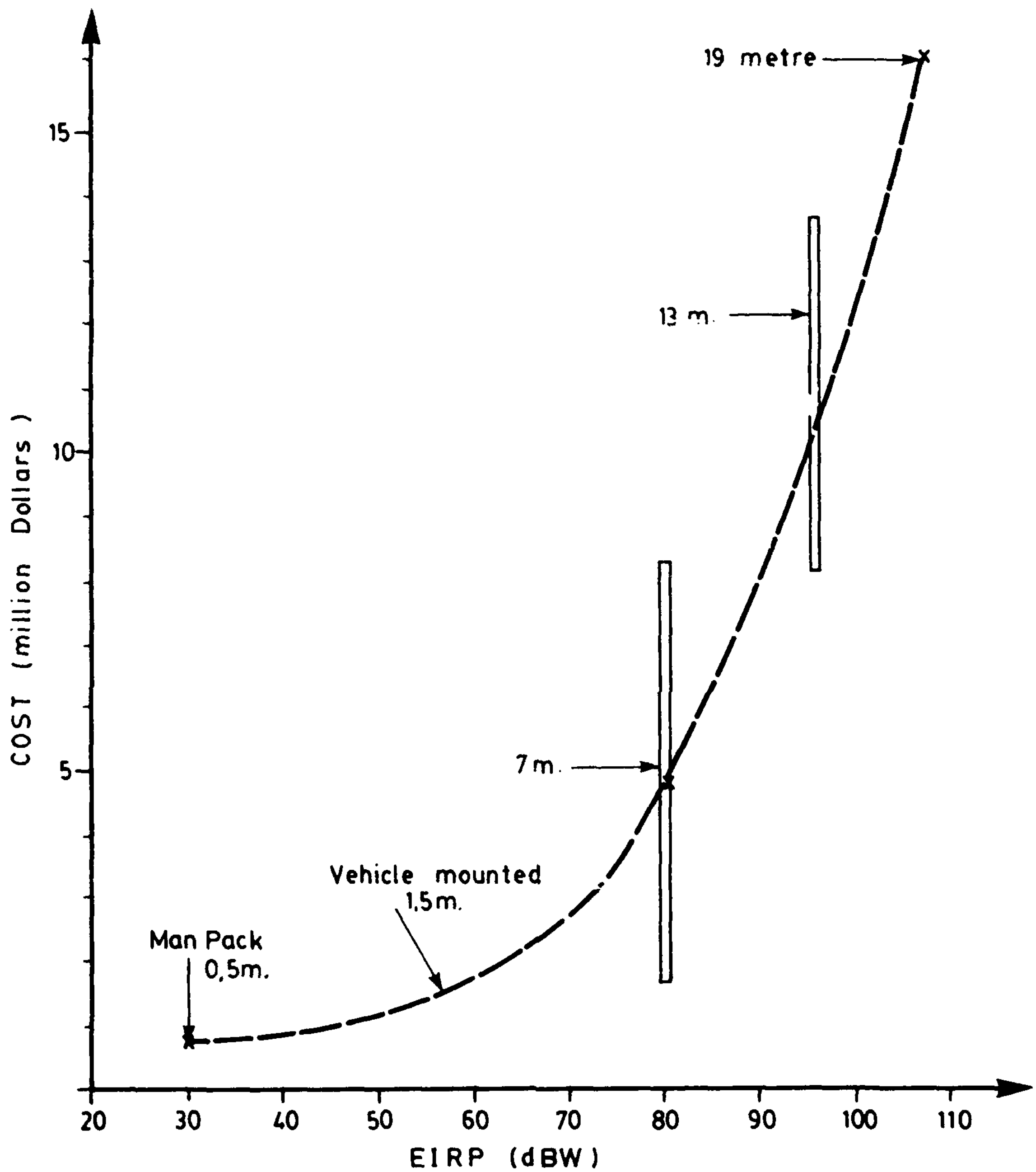
A practical solution to overcome the above-mentioned difficulties is the use of electronically steereable microstrip patch array antennas (Appendix 8.14 A). Such an antenna is mainly composed of a number of single array elements in combination with processor-controlled phaseand/or amplitude-shifters for beam forming and steering. Besides the advantages of a microstrip array antenna in general mentioned in the beginning, this electronically steerable antenna offers additional features such as:

- Adaptation to continuously changing reception situations for mobiles by electronic diagram forming and steering
- Fast antenna tracking especially suited for land mobile applications
- Blanking of unwanted signals (e.g. noise, multipath effects, satellite or terrestrial jammers) by electronic nulling.

The VSATs described above well-suited for military applications.The simple and rugged construction enables an easy transportation,i.e the VSATs are very well accommodated to changing battlefield conditions,further on a fast deployment of forces is possible

In combination with electronically steerable microstrip array antennas,an efficient mobile earth station with remarkable jammer resistance is feasible.

In view of the future development of sophisticated and efficient signal processors and multibeam antennas for satellites, an over-all system, tailored to military purpoeses can be established.



Fi.8.14.1 Current SHF satellite earth terminal costs

APPENDIX 8. 14A
ELECTRONICALLY STEERABLE ARRAY-ANTENNAS FOR MOBILE SATELLITE COMMUNICATIONS

1. SPECIAL REQUIREMENTS FOR LANDMOBILE SATELLITE COMMUNICATIONS

A major problem for satellite-based landmobile vehicle communications is the continously changing reception situation scenario with related large field-strength variations due to multipath effects (signal reflection, refraction and scattering), signal fading and shadowing by buildings or vegetation. Further on one has to deal with unwanted signals, e.g. those of other satellites or terrestrial sources, in the case of war-time possibly emitted by hostile jammers.

Civil landmobile communication terminals such as the Standard C terminal (INMARSAT) or the PRODAT terminal (ESA) are not unreser-vedly suited for military purposes: the quasi-omnidirectional conventional antenna system is too susceptible to interference caused by natural or artificial sources.

A practible solution to overcome the above-mentioned difficulties is the use of electronically steerable array-antennas. Some antenna systems of this special type are described in the following together with a discussion of the suitability and advantages for military landmobile communications as well as future prospects.

2. GENERAL PRINCIPLE OF ELECTRONICALLY STEERABLE ARRAY-ANTENNAS

An electronically steerable array-antenna is mainly composed of a number of single array elements in combination with processor- controlled phase-/ amplitude-shifters for beam forming and steering; the general function of such an adaptive antenna as in the case of the antenna described in chapter 3.1 is shown in Figure 1.

The receive signal vector x(t) is a mixture of the useful signal s(t), a noise signal n(t) and an eventual jammer signal i(t). By connecting the signal vector x(t) with complex weight factors w(t) followed by an addition and real part formation an output signal y(t) is obtained. The difference of this output signal y(t) and an appropriate quality signal d(t) is fed as differential signal e(t) to the signal-processor. In combination with an adaptive algorithm device the signal processor calculates the above-mentioned weight factors w(t) in an iterative closed-loop way by using the input signal information from x(t) and the differential signal e(t).

In the case of the "Steerable array-antenna for mobile digital sound reception" as described in chapter 3.1 the quality signal d(t) is equivalent to the bit error rate of the demodulated data stream, which is directly related to the S/N- ratio of the output signal y(t).

3. STEERABLE ARRAY-ANTENNAS FOR MOBILE RECEPTIONS

3.1 National Development

Main goal of this national funded project is the development of a quasi-flat array-antenna breadboard-model for mobile reception of satellite-broadcasted digital sound signals. Antenna diagram forming and tracking is performed by a signal-processor in combination with adjustable digital phase shifters.

This project itself started in 1987 and will be finished at the end of 1990. For a following phase the development of a more sophisticated array-antenna based on MMIC technology with much more smaller dimensions than this breadboard-model might be possible.

Main specifications of the antenna development are compiled in the following:

• Frequency Range	11.7-12.1 GHz (Ku-Band)
• Polarization	Circular
• Figure of Merit G/T	0 dB/K
• Reception Angle Range	21-35.5 in Elevation 0-360 in Azimuth
• Time Duration for first Acquisition	10 sec maximum
• Tracking velocity	50 deg per sec
• Number of Array Elements	about 256 (16x16 matrix)
• Antenna Gain	25 dBi in Main Direction
• Array Dimensions	about 300x300x50 mm (LxWxH)

3.2 Development at JPL

Two breadboard versions of a phased array-antenna have been independently developed by Bell Aerospace Corp and Teledyne Ryan Electronics through contracts monitored by the Propulsion Laboratory JPL.

The main performance data of the above-mentioned developments are:

• Frequency Range	1.5-1.6 GHz (L-Band)
• Polarization	Circular
• Reception Angle Range	20-60 in Elevation
• Time Duration for first Acquisition	10 sec
• Number of Array Elements	19
• Antenna Gain	8-12 dBic
• Array Dimensions	about 500-600 mm Diameter; about 20-30 mm height

4. USABILITY OF ELECTRONICALLY STEERABLE ARRAY-ANTENNAS FOR MILITARY MOBILE COMMUNICATIONS

The main advantages of this antenna type are:

- Adaption to continously changing reception situations for mobiles by electronic diagram forming and steering.
- Fast antenna tracking especially suited for landmobile applications.
- Blanking of unwanted signals (noise; multipath effects; satellite or terrestrial jammers) possible by electronic nulling.
- Small dimensions and easy integration into vehicle surface possible.

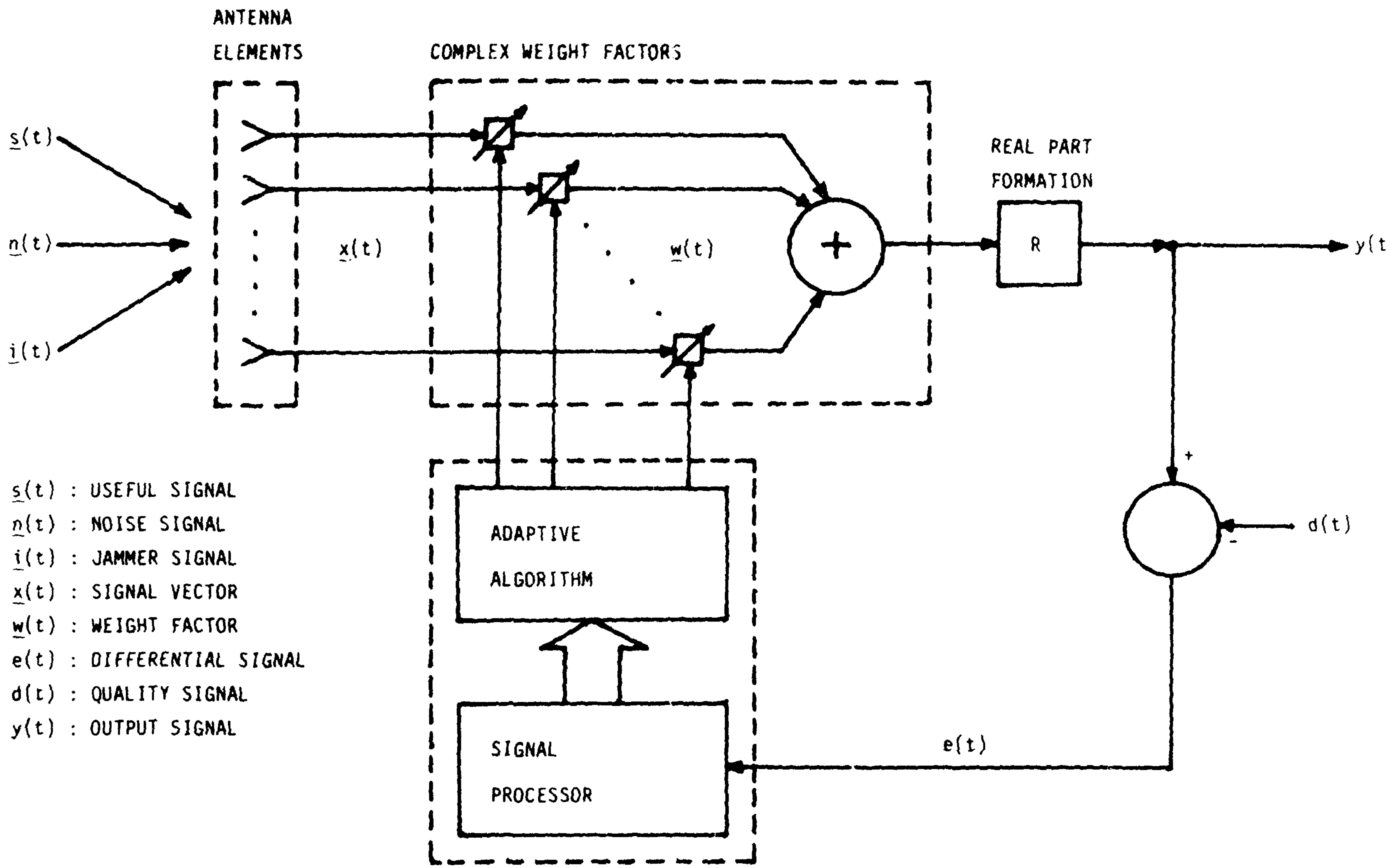

ANTENNA ELEMENTS
COMPLEX WEIGHT FACTORS
REAL PART FORMATION
R
s(t)
n(t)
i(t)
x(t)
w(t)
y(t)
d(t)
e(t)
+
-
ADAPTIVE ALGORITHM
SIGNAL PROCESSOR
s(t) : USEFUL SIGNAL
n(t) : NOISE SIGNAL
i(t) : JAMMER SIGNAL
x(t) : SIGNAL VECTOR
w(t) : WEIGHT FACTOR
e(t) : DIFFERENTIAL SIGNAL
d(t) : QUALITY SIGNAL
y(t) : OUTPUT SIGNAL

Figure 1

8.15 PACKET RADIO TECHNIQUES FOR SPACE USE

8.15.1 System Aspects

In packet switching,a general distinction, is made between two procedures.

- Message Switching. For the transmission, the message is split into blocks, if required, but at each intermediate node it is temporarily completed and in the correct sequence
- Packet Switching. The message is split into packets which are transmitted by sections, also via different paths. Each packet has a header with additional information (e.g. transmit and receive address, packet length, packet number, time mark, control bits for any error detection, etc.). The sequence of the packets must be restored at the target node or at the data equipment.

As for the satellite-based communication, only Digital Packet Switching is taken into consideration. The satellite itself can be *regarded* as a *receiving node which concentrates the* packet oriented data of the terminals for an efficient transmission. In view of the future development of sophisticated switching matrices and signal processors, an efficient routing flexibility is real-izable. Combined with a digital regenerative repeater, a decoupling of the up-and down-link is possible, further on a 100% signal restoration can be achieved.

Digital Packet Switching is a very efficient method for voice and/or data transfer; the main advantage are:

- Simple integration of different services into one network assignment to terminals being ready for transmission
- Utilization of transmission recesses of one terminal for engagement with packets of another one (e.g. TASI-system and/or Digital Speech Interpolation).
- Error detection and correction
- Implementation of vocoders for efficient bandwidth use
- Blocking-free communication
- Flexibility with respect to enlargement or temporal changes in density of communication traffic
- Simple implementation of broadcasting capability
- Easy use of efficient encoding methods

An essential drawback of Packet Switching, especially for satellite-based systems, is the occurrence of a considerable transit time mainly due to the long distance between terminal and satellite; *additional portions result from packet* generation, data processing, and storage. In case of voice transmission, a certain discipline has to be kept by the users. Another more general drawback is the additional need of channel capacity for the packet header, acquittance, status report, etc.

First Packet Switching experiments started in 1969 with the US- system ARPANET, a lot of other experiments and networks followed. Some of them together with the appropriate channel bit rate are complied below:

- ARPANET	50 kbits/s
- SATNET	64 kbits/s (use of INTELSAT-IV channels)
- PRNET	100-400 kbits/s
- LEXNET	1 Mbits/s
- Wideband (WB)-SATNET	772-3088 kbits/s (use of WESTAR-III channels)
- ETHERNET	10 Mbits/s

8.15.2 Conclusion

One of the essential tasks of a military communications system is the transmission of short messages: these can be data packets or voice packets. In view of the future development of sophisticated and efficient switching matrices and signal processors as well as regenerative repeaters for satellites, the use of intensive digital packet radio technology lends itself very favourably to this purpose.

8.16 RECOMMENDED PRINCIPLES FOR SELECTING NATO SPONSORED R&D PROJECTS

There is no doubt that the R&D work related to the technologies described above which are being carried out in national and multinational laboratories will, either wholly or partially, benefit the NATO future SATCOM systems. However, in the areas where the NATO needs are urgent or so different that the related technologies require adaption, then NATO should consider sponsoring R&D in these areas.

The Group believes that NATO should and can follow the examples of INTELSAT and INMARSAT and sponsor R&D work on a scale that is suited to NATO funding and interest in SATCOM. It is expected that funds would be rather limited and not adequate for the development of new technology for NATO SATCOM and a large scale. R&D work sponsored by NATO would be an efficient way of getting information about the results generated by ESA, NASA, and other national programs en a timely manner. As in INMARSAT and INTELSAT, potential suppliers to NATO are also expected to subsidize the R&D studies required by NATO and this would therefore be very cost-effective for NATO.

The current trend in civilian SATCOM which as away from large inter-city trunks and is toward thin route networks for business systems and mobile communications, is beneficial for NATO since this creates a need for the civil agencies to conduct detailed on-board signal processing studies. The civilian payload technology developed for this purpose could be transformed into AJ processing transponders at costs not necessarily unrealistic for NATO R&D funding. This applies to the AJ processor itself as well as as to the nulling (beamforming in civilian terms) uplink receive antenna.

It is shown in Section 8.2 that surface acoustic wave (SAW) devices can be used to provide some of the functions of an AJ processor and that digital technology can be used to enhance the AJ performance. Digital processing can, of course, in principle be used for the entire processing, but one has to consider the question of power consumption and other practical matters before deciding on the technique to use. AJ processing with a high degree of flexibility and to a large extent based on future civilian technology is therefore considered to be a good candidate for NATO sponsored R&D activity.

The uplink receive nulling antenna to combat uplink jamming is discusse in Section 8.3. It is quite conceivable that future civilian networks for mobile users may require similar capabilities in the form of phased arrays with frequency selective beam forming networks to control interference. Commitments of some R&D funds by NATO in this area may allow NATO to follow and influence the civilian efforts leading to lower costs, more competition and better and more flexible AJ antenna for everyone.

The civilian interest in the area of active antenna array technology will be great but not for the 45 GHz band MMIC technology will be nevertheless an important one for NATO to follow, but it may be difficult to accelarate the progress significantly due the considerable funding required in this area.

If laser inter-satellite links are of interest to NATO this area may also be a candidate for limited R&D funding which may be used to update the technology before entering into the specification phase of the satellite(s) to be acquired.

LASERCOM for submerged submarines which is discussed in Section 8.6 is a technology considered to be at present beyond the reach of NATO resources.

Technologies to support the spacecraft bus are also of interest to NATO for the purposes of reducing costs and improving performance and reliability, but this area is considered to be too general to suite NATO R&D funding objectives.

The areas of artificial intelligence and neural networks are being studied intensively in the member countries for different applications including communications (see Section 8.9). It is therefore expected that future NATO SATCOM would benefit from these efforts particularly in the field of autonomous control of spacecraft and in O&M generally. However, NATO should be ready to undertake or sponsor some adaption work in this field where the applications are NATO specific.

CHAPTER 9

NATIONAL SATCOM SYSTEMS AND DEVELOPMENT

9.1 INTRODUCTION

The meetings of the Group were held in the countries participating in the studies and at the laboratories and establishments of the members of the Group. At these occasions the opportunity was always taken to collect information from presentations given and the laboratories visited connected with satellite communications and related work areas.

A wealth of information obtained in this way combined with information provided by the members directly concerning R&D and planning work being undertaken in their countries on satcom systems forms the data base from which various sections of this report have drawn heavily. The present Chapter is no exception in this regard but includes brief desription of only those SATCOM systems, NATO and national, which are considered relevant to the work in hand and are not meant to be exhaustive.

9.2. UK SYSTEMS

9.2.1. Background

The UK became involved in satellite communications in 1966/7 by participating in the USA's IDCSP programme. The success of this venture prompted the UK to invest in a national military satellite programme, Skynet.

Skynet 1A was launched in November 1969 and placed in geostationary orbit. The success of Skynet 1 was, however, short-lived. The inorbit spare satellite, Skynet 1B, failed to achieve orbit when its apogee boost motor misfired. Shortly after that the life of the original 1A was brought to an end by transponder problems. Fortunately, the next phase of the programme, Skynet 2, was already underway.

Although Skynet 2A also suffered a launch failure, the second satellite was successfully launched in November 1974 and provided service for more than a decade, even though the ability to maintain its station was eventually lost.

These years saw a steady reduction in the UK military presence worldwide in favour of a concentration of UK forces in Europe and the Atlantic. The Skynet 3 system was cancelled and as a continuation of the long-standing US/UK cooperation in military satellite communications the UK utilized access to US Defence Satellite Communications System DSCS II satillites. The reduction in perceived strategic communications led to a reduced Army and RAF satellite communitcations requirement. However, it was the Navy who began to realize the full value that satellite communications would have for its ships operating throughout the oceans.

The Falklands War gave sharp and timely emphasis to the need for reliable long-range communications. Conditions for HF communications were poor but, fortunately, Skynet 2B was towards the western extremity for its drift orbit. The Navy was able to speak directly and securely from the South Atlantic to the Joint HQ at Northwood through their SCOT ship-borne satellite terminals. Also the Army, once ashore, used tactical UK/TSC 502 satellite terminals to communicate with the fleet as well as with UK. The need for UK military satellite communications was demonstrated convincingly.

9.2.2 Skynet 4

In the early 1980s the Navy were convinced of their future dependence on satellite communications and they took the lead in creating the requirement for the Skynet 4 satellites. Skynet 4 is a two-stage programme for the Space Segment with Stage 1 providing three satellites.

The Space Shuttle was scheduled to launch 4A and 4B in 1986 and 1987 respectively, but following the Challenger accident alternative launch vehicles had to be selected. The first Skynet 4 satellite was successfully launched using Ariane 4 in December 1988 followed by the second in December 1989 using a Titan launch vehicle. Launch of the final Stage 1 satellite, again using Ariane 4, is expected in September 1990. Further satellites will be required some 5 to 7 years from the start of the programme as the Stage 1 satellites approach their end-of-life.

Skynet 4 is a compact satellite suited to UK military needs. The basic satellite vehicle is a continuation of the British civil European Communications Satellite (ECS) series. Since the formation of the European Space Agency (ESA) in 1975, the British space industry has concentrated on civilian satellite development. An Orbital Test Satellite was successfully launched in May 1978 and became ESA's first communications satellite. Refinement of the design led to the Maritime European Communications Satellites (MARECS) which started life in 1981 and the ECS which started in 1983. Skynet 4 therefore has a mature and well-tried satellite platform. The prime contractor is British Aerospace who, with Marconi Space Systmes, lead a wide consortium of European partners, providing many of the subsystems.

On station the satellites are 3-axis stabilized, they do not spin to maintain pointing stability but use internal gyroscopes. This allows the body to be static and long solar array wings can be used to generate more power than is available to a spinning satellite. With a wingspan of around 16 metres, the arrays provide about 1.2 kW to the satellite for its sub-systems and for the communications payload. The initial on-station mass of the satellite is around 800 kg.

The Skynet 4 payload includes both SHF and UHF transponders. These provide a total of 120 W output power at SHF and 60 W at UHF. Multiple SHF antennas provide earth cover and area coverage beams, including a spot- beam over Central Europe. The SHF transponders have a total bandwidth of 340 MHz. The payload also includes an experimental EHF receiver operating in the 43.5-45.5 GHz band.

From its position above the equator at 1° W, Skynet 4 is able to 'see' about a third of the Earth's surface including most of the Atlantic and Europe. The Navy will be the major user of the earth-cover antenna and of the two UHF channels for very long-distance maritime communications. The Army have a special interest in using portable, relatively low-power terminals in Central Europe and the efficient spot-beam will be especially useful to them. The RAF will also use this for UK and RAF Germany communicatinos and will use the wider beams for Cyprus and for maritime air operations. The experimental EHF receiver will permit furtner research into future use of these upper bands which offer very high bandwidth and, therefore, information capacity, together with extremely good ECCM characteristics.

9.2.3. Ground Segment

The main ground installation to operate, control and manage communications through the Skynet 4 satellites is RAF Oakhanger. To meet the much increased command and control communications requirements for the UK military satellite

communications system, RAF Oakhanger has been subject to an extensive enhancement programme The centre of this programme has been the design and construction of a new communications and control building New tracking, telemetry and control equipment allow the operators to monitor the status of satellite sub-systems, switch between redundant systems as necessary and maintain the satellites in the desired orbital locations

The communications sub-system of the satellite is also controlled here, and RAF Oakhanger will continue to act as the hub of the communications network using the satellites The major features outside the new building are the new ground antennas of the Plessey UK/646 and Marconi UK/TSC 648 type

Through the Skynet 4 satellites, RAF Oakhanger provides communications between fixed terminals such as those in Cyprus and Germany and mobile terminals operated by all three Services

Tactical terminals that use Skynet 4 include UK/VSC 501 Landrover-based systems such as those for RAF Germany and for the mobile Tactical Communications Wing at RAF Brize Norton The latter will be used as communication rear links for contingency and out-of-area operations Tactical terminals developed for Skynet 4 include small "manpack" UK/PSC 504 terminals These terminals collapse to form a backpack which can be carried by one man

The installation of satellite communications terminals into aircraft is being investigated but there are some special problems-mainly finding space for the antenna and providing a stabilization system to keep the antenna tracking the satellite Despite these a terminal suitable for maritime patrol aircraft such as Nimrod has been developed by Marconi Space Systems

The Navy are in the process of equipping all their major warships with SCOT SHF SATCOM terminals They will also use the Skynet 4 UHF channels for communication with submarines

A current project of note is the integration of SATCOM into the British Army's tactical area communication system, PTARMIGAN

9.2.4. Future plans

Skynet 4 Stage 2 may include additional features to extend the flexibility of the system These include a steerable spot-beam antenna, commandable UHF channel frequencies and a small-scale 44 GHz 20 GHz EHF repeater possibly involving on-board processing

Planning for a follow-on space segment is still at an early stage, but it is thought that the emphasis will be on the introduction of an operational EHF SATCOM system, while at the same time providing continued support for the existing SHF ground segment

A major enhancement of the UK military SATCOM system involves the diversification of the ground segment through the addition of alternative "anchor" stations at well-separated locations in the UK that can be used in the event of the failure of RAF Oakhanger This project is due to be completed by the end of 1992

9.3. SYSTEMS AND DEVELOPMENTS IN THE USA

9.3.1. The NASA Advanced Communicatiouns Technology Program (ACTS)

As a direct result of a 1978 Presidential order, NASA assumed the responsibility for developing new technology to increase the capacity of future domestic satellite systems The ACTS Program is the vehicle for reaching that goal

The program encompasses the satellite its communications ground terminals and its control segment The ACTS users comprise a consortium of academic industry and government agencies whose interests are to understand and evaluate the Ka-band technologies The ACTS Program incorporates several examples of new technologies, such as the multi-beam antenna the IF switch matrix, the baseband processor, rain-fade compensation, signal processing low-noise receivers and serial minimum shift keyed (SMSK) modems

The multiple beam antenna on the ACTS satellite (Figure 9 3 1) provides both fixed spot beams and scanning beams The ACTS communication subsystem (Figure 9 3 2) operates simultaneously in two different modes the low burst rate mode and high burst rate mode, which satisfy separate sets of communications needs The high burst rate mode demonstrates high rate trunk communication links between large cities It uses the three fixed beams, one points at Cleveland Ohio where NASA's ACTS network control station is located. The other two point to Atlanta Georgia and Tampa Florida. Also, one of the scanned beams which are normally used for the low burst rate mode may be fixed-pointed at other sites in the United States and substituted for the Atlanta or Tampa beam for use as a high burst rate beam The high rate transmissions are TDMA bursts at 220 Mbps using either QPSK or SMSK modulation Power control is used in both ground terminals and the satellite to compensate for atmospheric attenuation variations

Figure 9 3 1 - Figure 9 3 2

The purpose of the low burst rate mode is to demonstrate flexible communications for widely distributed users with smaller capacity requirements. It normally uses two scanned beams each covering one of two adjacent areas in the northeast United States. In addition, the two scanned beams can provide spot coverage to 13 other cities where ACTS experimenters are located The low rate mode supports two 110 Mbps or eight 27 5 Mbps TDMA bursts in each uplink beam The bursts within a beam are separated by FDMA In the satellite all bursts are routed to the baseband processor which demodulates sorts routes, reformats, and remodulates the signals into one 220-Mbps TDM data stream per downlink beam SMSK modulation is used on both uplinks and downlink The processor can also support error correction coding on selected uplinks and downlinks to counter the effects of rain attenuation

The ACTS program is basically experimental In addition to NASA's experiments to verify the newly developed technology the consortium of independent experimenters has proposed a variety of experiments Examples include one meter dishes communicating at 6 9 Mbps high burst rate mode communications with small dishes at reduced rates communications to and from mobile platforms and propagation studies in many U S locations

The experimenter network will consist of a number of hub terminals with each terminal networked to various user computers The hub terminal (micro-terminal) controls the TDMA burst plan coding, modulation and microwave receive and transmit functions These micro-hub terminals cannot operate as stand-alone units; support is provided by both the micron control center and the users

9.3.2. FLTSATCOM AND AFSATCOM

FLTSATCOM satellites are the spaceborne portion of worldwide Department of Defense communication system that provides robust anti-jam communications links between aircraft, submarines, ground stations, strategic force elements and the presidential command networks Each satellite is deployed in a geostationary orbit where it provides visible-earth coverage over 23 channels in the 240- to 400- MHz frequency band and at SHF The FLTSATCOM satellites are built by TRW, Inc

Tacsat and LES-6 were experimental satellites that demonstrated

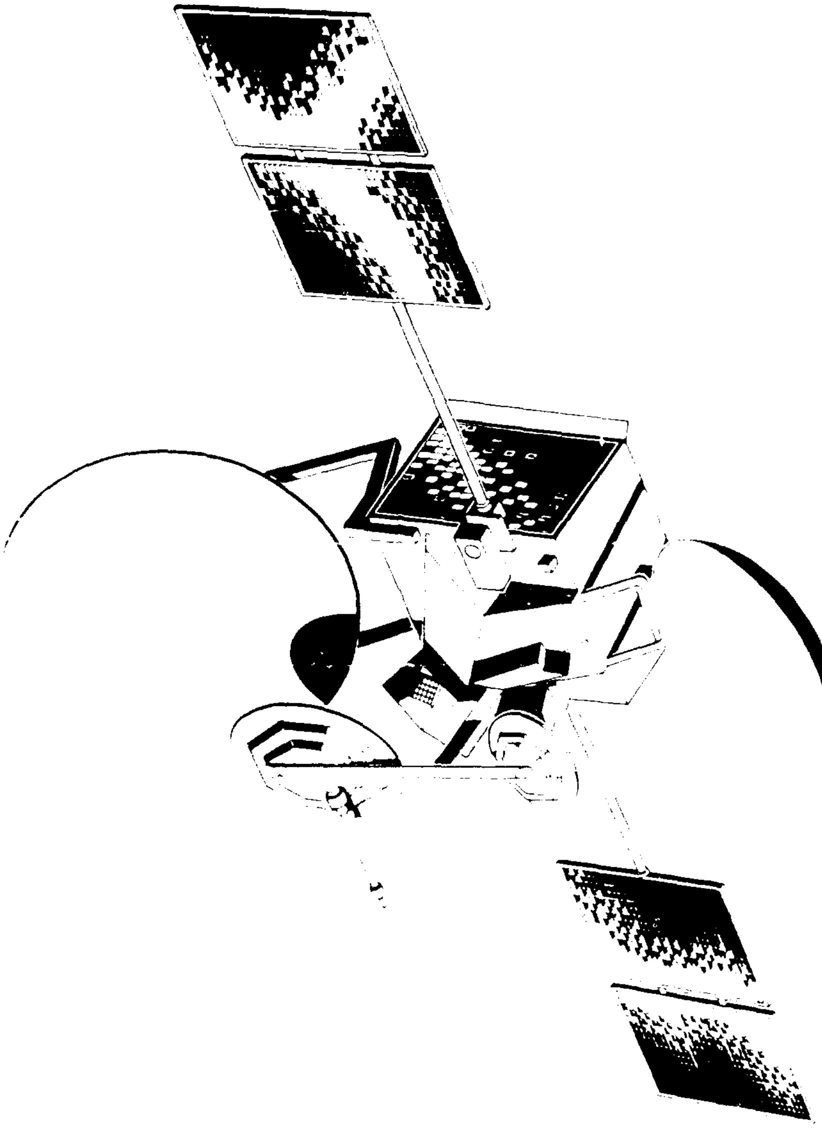


Fig 9.3.1 The NASA ACTS satellite

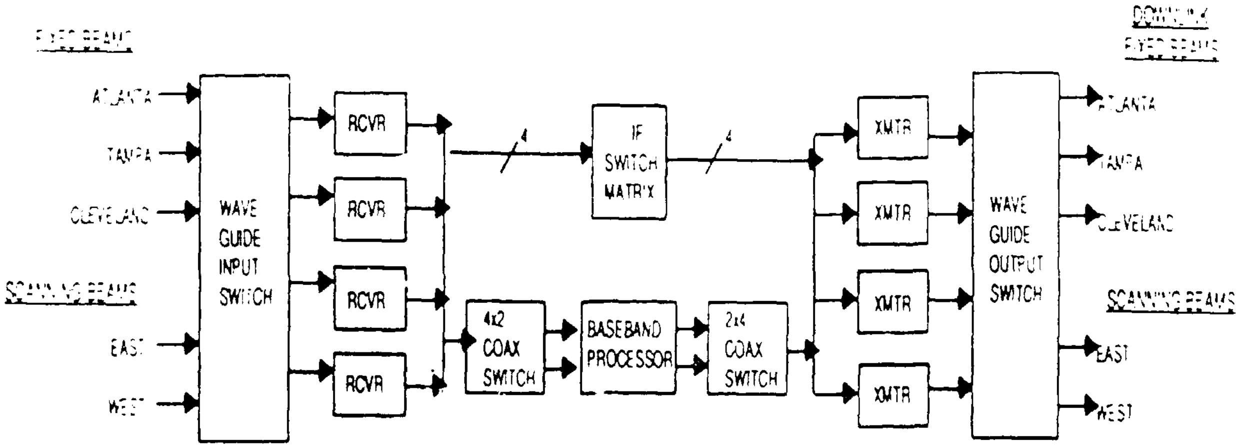


Fig.9.3.2 The ACTS SATCOM system

the first UHF communications satellite links with mobile terminals in the late nineteen sixties and early seventies. These satellites were used for numerous tests and provided a limited operational capacity. Tacsat ceased to operate at the end of 1972, and LES-6, although considerably degraded, is still operational in a stand-by status.

The FLTSATCOM system is the first operational (rather than experimental) system for mobile tactical military users. FLTSATCOM serves Navy ships, aircraft, and shore stations. AFSATCOM serves Air Force strategic aircraft, airborne command posts, and ground terminals. The two systems share FLTSATCOM satellites in synchronous equatorial orbits. The Air Force also has communications packages on several other satellites in highly inclinated orbits to provide coverage of the north polar region.

FLTSATCOM satellite numbers one through six have a hexagonal body composed of two modules (Figure 9.B.3.). The service module (to the rear) contains the attitude control power, and Figure 9.3.3 TT&C subsystems as well as the apogee motor. The forward module contains the communication subsystem. The antenna for UHF transmisssions is a 16-foot diameter paraboloid with a solid center section and an outer surface of mesh attached to ribs. The mesh is deployed after the satellite is injected into synchronous orbit. The UHF receiving antenna is a separate helix about one foot in diameter and eleven feet long. The third antenna is an X-band horn for reception of the fleet broadcast uplink and for transmission of a beacon.

The satellite has four types of communication channels. The Navy uses one fleet broadcast channel and nine, 25-kHz bandwidth fleet relay channels. The Air Force uses twelve narrowband (5 kHz each) channels and one wideband (500 kHz) channel. All links, except the fleet broadcast uplink, are in the 240to 400- MHz band with the downlinks at the lower part of the band. The fleet broadcast uplink frequency is at 8 GHz. Either processing or non-processing receivers may be used with the fleet broadcast and some of the Air Force narrowband uplinks. The processsing receivers provide an anti-jam capabilitiy

The satellite has twelve transmitters, one for each of the Navy channels, one for the Air Force narrowband channels, and one for the Air Force wideband channel. A UHF command channel allows operational control of the communications package and for limited redundancy switching.

The fleet broadcast channel carries fifteen teletype and one synchronization channel for a total data rate of 1200 bps. The 25 kHz fleet relay channels are normally used in conjunction with TDMA automated demand assigned multiple access (DAMA) techniques. The TDMA DAMA format uses burst rates between 9.6 and 32.0 kbps. Each narrowband Air Force channel carries single 75-bps link. The wideband channel carries either multiple FDMA links at 75 bps or a single higher rate link.

For polar coverage, AFSATCOM communications packages are placed on other DoD satellites. These packages have capabilities similar to those of the twelve narrowband Air Force channels on the FLTSATCOM satellites. In addition, AFSATCOM includes a single-channel transponder with anti-jam capabilities carried on each of the DSCS III satellites.

9.3.3. FLTSATCOM EHF PACKAGE

On FLTSAT satellite numbers 7 and 8, the MIT Lincoln Laboratory has provided an EHF package (Figure 9.3.4.) which is a forerunner to MILSTAR. FLTSAT 7 was launched in 1986 and FLTSAT 8 was launched in 1989.

The objective of the FEPs on FLTSATCOM is to test the operational capabilities of developmental EHF MILSTAR terminals and to provide an early operating capability for certain key functions of the MILSTAR system. Eventually, ships, submarines, fixed-shore installations, aircraft, and ground mobile forces will be equipped with EHF SATCOM terminals.

The FEP is housed in a third hexagonal module added to the rear of the FLTSAT spacecraft. The FEP antenna looks out through a hole cut in the solid portion of the UHF transmit antenna. This hole is covered with a window which is transparent at EHF frequencies but reflective at UHF frequencies and which provides Figure 9.3.4 a thermal barrier over the opening.

Figure 9.3.5. is a simplified functional block diagram of the FEP communications system. The FEP operates at EHF frequencies of approximately 20 GHz on the downlink and approximately 44 GHz on the uplink. It has two antenna beams: a 5 degree spot beam steerable by ground command and an earth coverage beam. The 5 degree spot beam is generated by a single steered dish with a dual frequency feed for the uplink and dowlink. The earth coverage beam is implemented by separate horn antennas for the transmit and receive frequencies.

Uplink signals are frequency hopped; the basic uplink signal format is frequency-division multiple access. Multiple users at a terminal and overhead control functions are time-division multiplexed within each frequency channel (FDMA/TDM). Multiple-frequency shift keying (MFSK) modulation is used. Downlink signals are time-division multiplexed (TDM) and frequency hopped. Modulation is either differential phase-shift keying (DPSK) or MFSK. There is a single downlink carrier which can be switched between the two antenna beams on a hop-by-hop basis by the ferrite antenna switch.

A key feature of the FEP communications system is the application of surface-acoustic-wave (SAW) chirp/Fourier-transformation techniques to demodulate--with minimum demand for dc input power--the MFSK signals received simultaneously in many frequency bins. Carrying out the demodulation function with the digital technology available at that time would have been prohibitively costly in terms of power. Because of subsequent rapid advances in VLSI technology, that may no longer be the case today.

A second key feature in the FEP is an autonomous, computer-based resource controller that autonomously sets up data channels at different data rates, via differenet antenna beams to support individual user-communications needs. User-service requests are received via an order wire from each user terminal's computer by FEPs on-board access controller. The controller in turn sets up the requested services and informs the involved user terminals'computers of its actions via a downlink order wire. Figure 9.3.5. gives an indication of the protocol messages that flow between the FEP and two terminals during the establishment of a communications link. Although the computer-to-computer dialogues between the FEP and the user's terminals are complex, the required human/machine interactions are user-friendly and can be easily performed by military personnel.

Once a channel has been set up, the FEP's high speed digital signal processing logic converts uplink message formats to downlink message formats and retransmits user data via either or both of the FEP's two antenna beams. Because this processing is done at baseband in digital logic, the switching of individual communications circuits within the satellite is straightforward to implement. This capability together with the autonomous resource controller makes FEP a true "switchboard in the sky", a concept that has long been in the minds of satellite system architects.

A fuller account of FEP communication system is given in Appendix 9 A.

9.3.4. MILSTAR

MILSTAR is the cornerstone of the U.S. government's program to

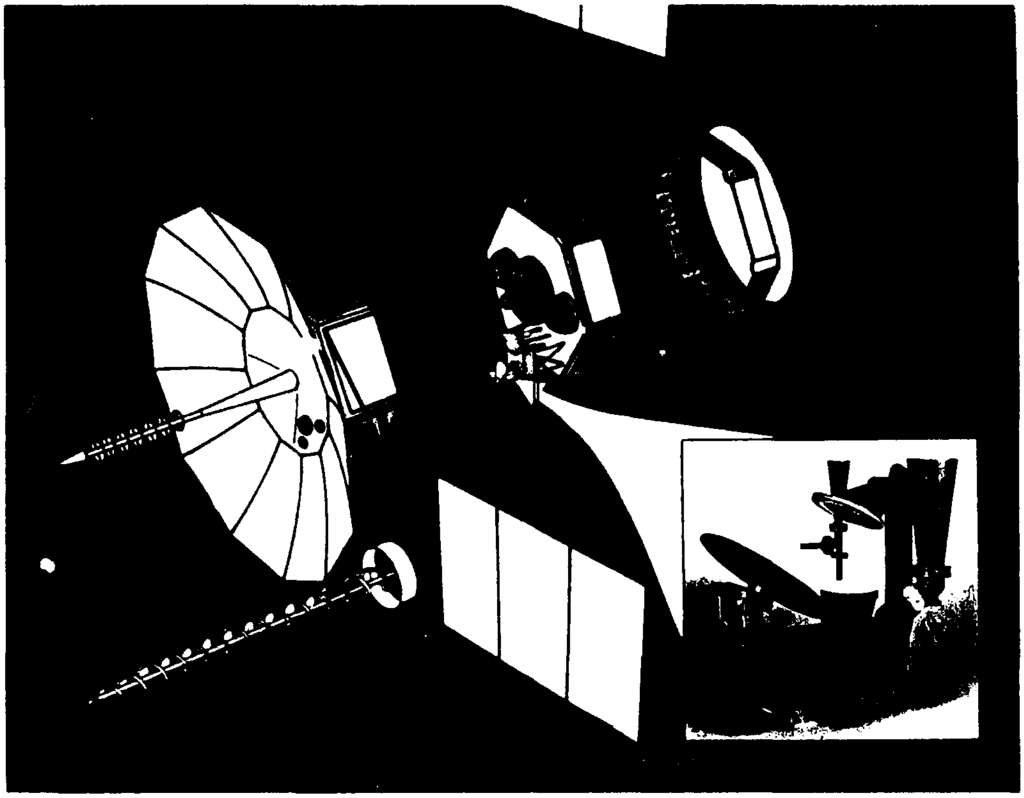
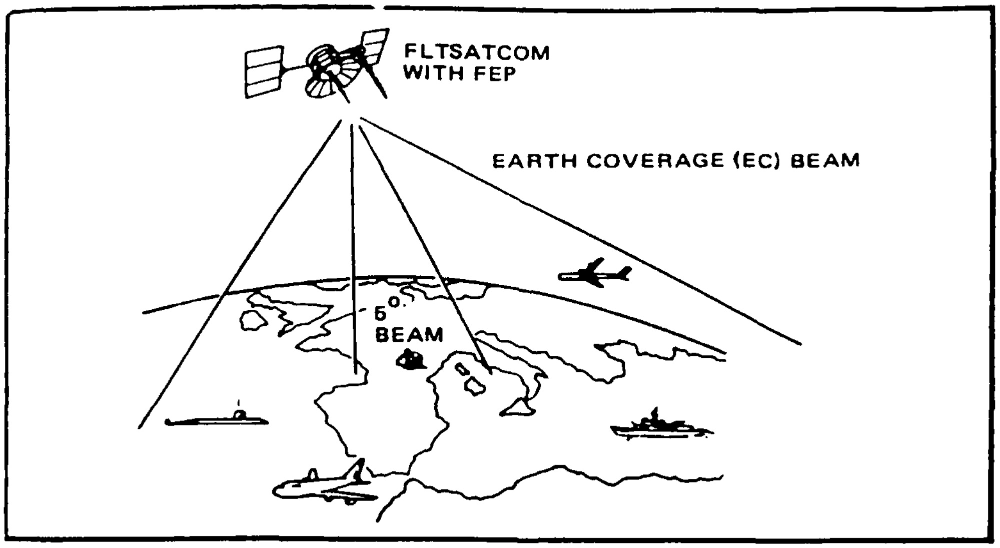


Fig.9.3.3 The FLTSAT spacecraft



- MILSTAR-compatible
- 5 deg mechanically scanned beam
- Earth-coverage beam
- Max. capacity 26 voice channels
- On-board access control
- Allows OT&E of MILSTAR terminals
- 20-watt redundant TWTAs
- Submarine report-back
- Electronics/antennas, weight 245 lb
- Bus power under 305 watts

Fig 9.3.4 The FLTSATCOM EHF Package (FEP)

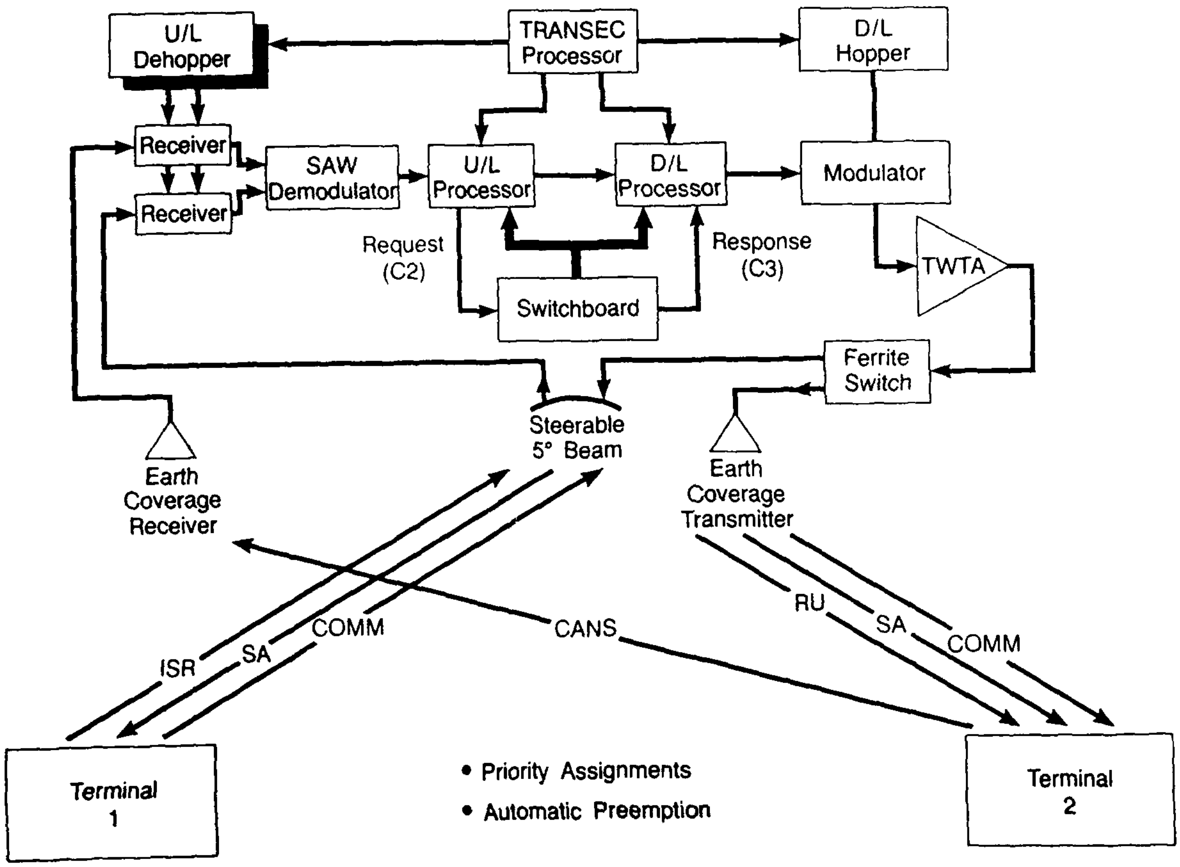


Fig.9.3.5 FEP Communications block diagram

make satellite command and control links as survivable as the units they control. MILSTAR serves both strategic and tactical users with the emphasis being on the strategic users. The capacity of the MILSTAR system is relatively modest by commercial standards, however, the low capacity links are extremely robust in the face of jamming and physical attack.

MILSTAR includes many of the features which FEP has, but has additional capabilities as well. It includes advanced EHF antenna technologies for AJ uplinks steerable downlink beams. It carries 60 GHz cross-orbit links for direct satellite-to-satellite data transfer. In addition to the EHF frequencies, which are identical to FEPs, MILSTAR covers the UHF spectrum for backward compatibility with existing ground terminals.

The MILSTAR constellation will eventually consist of ten satellites launched into high earth orbit by Titan launch vehicles. The Lockheed Corporation is the prime contractor. TRW provides the EHF equipment and E-Systems provides the UHF equipment.

EHF terminals for MILSTAR are under development by several contractors for the three services. All terminals are built to a common set of waveform and protocol standards. Terminal developments are well along and all terminals have been tested with the FEPs to verify interoperability and compatibility.

The MILSTAR program has encountered significant cost and schedule difficulties. The present status is that the first of the MILSTAR satellites is undergoing integration and test on the ground. Three of the eventual ten satellites have been funded.

9.3.5. LIGHTSAT

This is not as much a national program as it is a philosophy or a movement. This phlosophy emerged several years ago and has undergone several evolutions since then. It has been known under various names such as CheapSat, SmallSat, and Multiple Small Satellite Program (MSSP). Over the years DARPA has been the primary motivating force for the LightSat appreach, although the phlosophy is now rapidly gaining adherents in the military services and the commercial sector as well. The United States Congress recently increased DARPA's funding request for their LightSat Program.

This phlosophy originally sprang from the work on packet switched data networks at DARPA (ARPANET). One early concept was a worldwide packet switched network consisting of several hundred (220 was one number considered) satellites interconnected with each other by crosslinks and with ground nodes by uplinks and downlinks. The satellites would be very low in cost, hence in low orbit. The system would be robust beacuse of the large number of alternate paths that would exist between any two fixed ground nodes. It would be survivable because of the proliferation of satellites and finally it was hoped that service rivaling that of existing satellite systems could be provided at lower cost.

The main technical issues were: (1) The complexity of algorithms needed for the dynamic-node network of satellites in different orbits, (2) the difficulty in predicting exactly how well such a system would perform, and (3) the cost could not be made competitive with higher orbit systems, which required many fewer, albeit more expensive, spacecraft to obtain equivalent coverage.

Missions other than conventional space communications have also been considered as applications for the LightSat phlosophy. Of these, store and forward communications and signals intercept appear particularly attractive. These are missions which can be carried out with simple payloads and for which the limited coverage of low orbits is of little concern. Other applications might deal with the reconstitution of essential communications and reconnaissance capabilities after a nuclear strike.

The LightSat philosophy today encompasses several related efforts. First there is the DARPA program of the same name. This includes efforts to develop low cost launch vehicles such as the Pegasus which is being designed by Orbital Sciences Corporation to be launched from an aircraft into low earth orbit. The Navy is pursuing LightSat applications under the name SPINSAT and will test a Pegasus launched payload this year.

The Air Force has a similar effort going under the name 'Reserves'. The Army is looking at a number of applications as well. The LightSat concept was recently embraced as a part of the future military communications satellite architecture for the United States.

It is foreseen that the LightSat approach will make a significant impact on the future utilization of space in both commercial and military applications. It is unlikely, however, that LightSat will completely supplant current satellite systems or their lineal descendants. LightSats are seen as an augmentation and expansion of our present uses of space but not as a total revolution.

9.3.6. DSCS

The principal MILSATCOM system of the United States is the Defense Satellite Communication System (DSCS) which operates in the SHF military band (7 GHz down, 8 GHz up). It grew out of the IDCSP, a system of low orbiting unoriented SHF satellites that, in the middle 1960's became the nation's first operational MILSATCOM system. IDCSD was redesignated as DSCS I. The DSCS II satellites were spin stabilized in geosynchronous orbit and had de-spun and independently steered high gain antennas as well as earth coverage antennas. They have been in service since 1971. The last DSCS II was launched in 1982, along with the first of the new generation DSCS III's.

The DSCS III satellites are technologically very advanced. They are three axis stabilized and carried electronically switched SHF multiple beam antennas of advanced design. The uplink antenna (with 61 beams) can form variable gain patterns ranging from earth coverage to spot beams to patterns with adaptively formed nulls on jammers. The downlink antennas have 19 independently switched beams. DSCS III provides a wide range of services from very high data rate point-to-point links to low data rate links to ground mobile force. SHF man-portable terminals are now being developed for use with the DSCS system.

DSCS III satellites sufficient to last through the year 2000 have been procured. It appears that, with the next generation satellite, DSCS will transition to the military EHF bands (20 GHz down, 44 GHz up). The EHF follow-on to DSCS III will offer a range of new capabilities and services. It will serve very high data rate users with transponders, which will be frequency hopped for AJ protection. It will also offer EHF on-board processed channels for extremely robust low and medium data rate users.

The antennas will include various beamwidths ranging from earth coverage to very high gain agile spot beams. Antenna nulling will be included as it is on DSCS III, to provide additional jamming protection. At present the DSCS follow-on program is still in the early phases of definition. Several contractor teams have produced conceptual designs of all segments of the system--space, control and terminal segments. Supporting technology developments are underway in several quarters.

9.3.7. Strategic Defence Initiative (SDI)

SDI program could have major long-term implications for future civil and military space systems and technology developments as outlined below.

i) Much higher levels of on-board data and info processing

a) Gigabit range

b) Large scale internetting of sattellites and ground elements

c) With extreme reliability of both hard and software.

d) Efficient creation of massive and easily-validated software packages

e) Development and implementation of very large distributed systems including use of artificial intelligence.

All these are required for robust and survivable C3I.

ii) Great stress on survivability techniques in te... s of nuclear conflicts and attacks, and directed energy weapons (protection against these effects and antisatellite (ASAT) capability)

iii) Advanced space-based sensors including long-wave infrared sensors, mm-wave radars and laser imagers (All these are required for having exact information about a large number (10.000) of objects in space in a short time)

iv) High-power directed energy and kinetic energy weaponary of SDI would require supporting technology eg.

a) La.ge/Rugged optics

b) Precise stable pointing and tracking

c) Advanced propulsion techniques for small payloads

d) Large amounts of space power and energy

v) Complex space structures (construction, control, survivability)

vi) On-orbit maintenance and repairs

a) including autonomy, automation and robotics

b) aids to the detection and negation of potential threats (orbital debris, space mines, orbital bombardment satellites).

c) Capability to determine quickly the difference between an attack and an accident and to assess the damage involved along with the need for repair and rescue.

vii) Higher rate, heavier-lift launch capability and for greately improved cost effectiveness of launch systems and operations on Earth and in orbit.

viii) To invent durable techniques for achieving the above.

9.4 CANADIAN SYSTEMS

9.4.1. Background

In Canada, there has been a very successful history of accomplishment in civilian satcom. However, on the military side the effort to date has been very low level and mostly on the R&D side. For example, field trials of the government built, UHF, Canadian Shipborne Terminal during 1980 were well received by the Navy. More recently two 12/14 GHz terminals have been produced for ongoing field trials with the Army. They use a 25kHz channel on a civilian ANIK satellite. One terminal is vehicle mounted for rapid deployment and has been flown to numerous remote locations in a Hercules. In order to make future use of civilian satellites, encryption is being installed on the command link for the fortcoming ANIK-E satellites.

The largest current milsatcom project within the Department of National Defence (DND) is the D6470 EHF Satcom Project. It is a six year, $48M (CDN) project. The objectives of the project are to develop an extended technology base within DND, other government agencies and Canadian industry to support and reduce the risk associated with the possible future acquisition of an EHF military SATCOM system. A major portion of the effort is on the FASSET contract which will be described below. Other D6470 work includes technology development in higher risk areas such as nulling multiple beam antennas, and inhouse development at Defence Research Establishment Ottawa of a brassboard payload and two terminals.

9.4.2. FASSET

FASSET is an acronym for "Functional Advanced Development Model EHF Satellite Communication System for Experimentation and Test" The $28M Phase 2 contract started in March 1990. The FASSET will be comprised of; 1) one functional ADM satellite communications payload, 2) one functional ADM ground terminal, and 3) one ruggedized functional ADM ground terminal. When these items are delivered, they will be used for an extensive system and technology evaluation program to be performed primarily in government laboratories.

The FASSET system will be capable of providing either low data rate (LDR) service (user data rates up to 19.2 kb/s) or medium data rate (MDR) service (user data rates up to 1.544 Mb/s) as fully independent modes of operation. The two modes are not required simultaneously. All transmitters need only operate with a single carrier with the EIRP selected to develop a receiver SNR over terrestrial line-of-sight links that would be representative of links between a satellite ground terminal and a satellite in a selected orbit. The system will be capable of simulating geostationary and non-stationary orbits. Rather than introduce the effects of range and range-rate by means of a propagation simulator, provision will be made within the payload ADM to simulate the appropriate signal delay and doppler shift.

ECM protection will be provided by the use of frequency hopping on both the 44-GHz uplink and the 20-GHz downlink. The FASSET payload will be equipped with simple horn antennas. Nulling multiple beam antennas are being developed separately from FASSET but could be later combined with the FASSET payload.

One of the terminals will have a tracking antenna and will be transportable and ruggedized for operation and relocation within the Ottawa area. The second terminal will have a fixed antenna and will be constructed for a laboratory environment. A simplified block diagram of the terminal, showing both the LDR and MDR mode, is provided in Fig. 9.4.1.

These terminals will be generic EHF satcom terminals providing voice and data service rather than terminals designed for a specific application.

A simplified block diagram of the FASSET payload is provided in Fig. 9.4.2. The payload will be fabricated for a bening environment but provision will be made for positioning the antennas for communication with transportable terminals at some line-of-sight external location.

After acceptance of the FASSET system, it will be tested by government researchers to investigate the synchronization and communication performance in the presence of system noise and jamming. The various options in signal processing will be evaluated for effectiveness. Each sub-system will be electronically stressed to determine it's vulnerability to jamming.

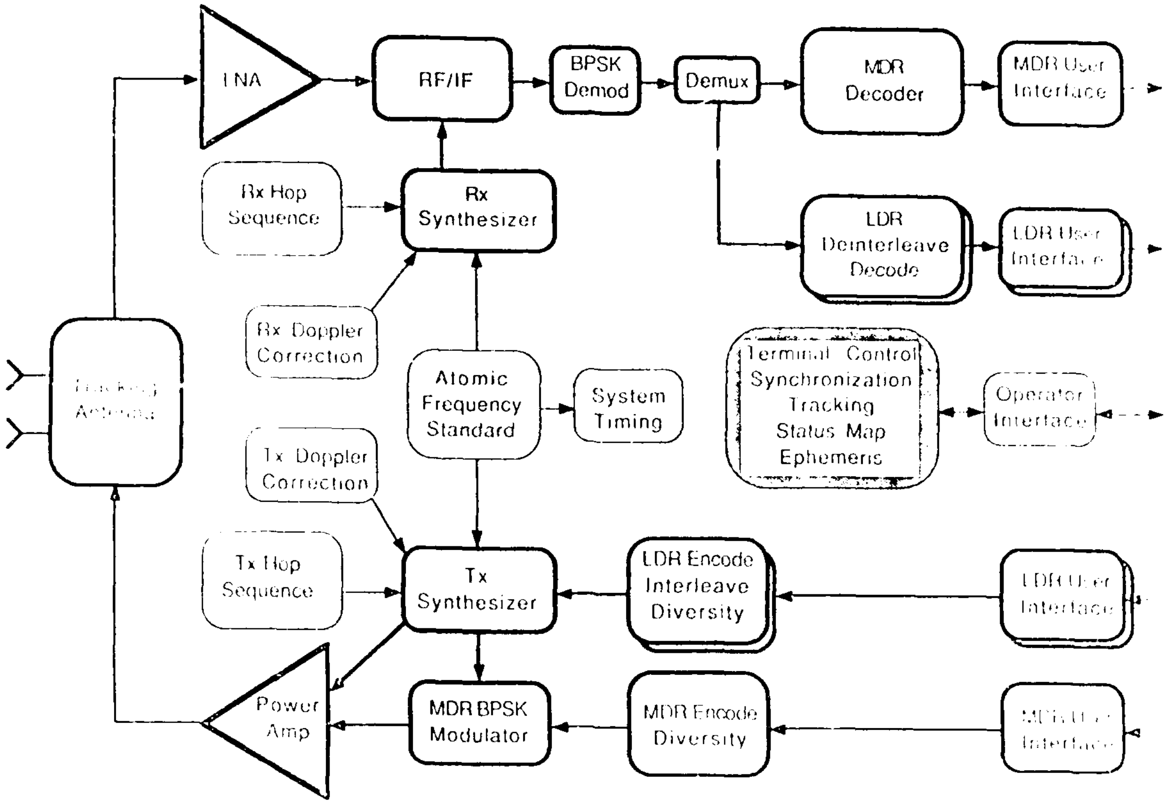


Fig.9.4.1 FASSET LDR/MDR ground terminal

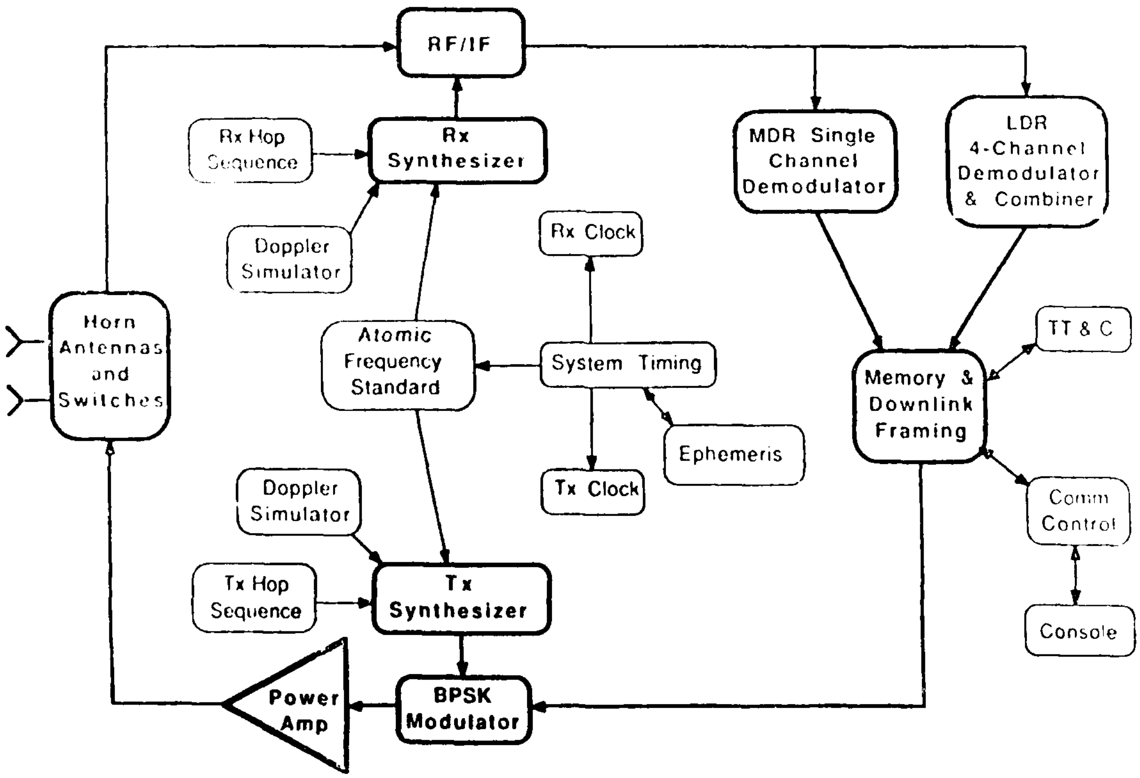


Fig.9.4.2 FASSET communications payload

9.5 FRENCH SYSTEMS AND DEVELOPMENTS

9.5.1 Present System - Syracuse I

The french military satellite communication network currently in operation uses the Telecom I multipurpose satellites, developed and operated under the leadership of the French PTT, General Directorate of Communications (D.G.T) now called France-Telecom.

The military payload of Telecom 1 (SYRACYSA I) includes two global coverage X band (7-8 GHz) transponders of 40 MHz bandwidth and equipped with 20 watts TWTs.

The system has been in operation since the first quarter of 1985 following the launch of the first satellite, Telecom 1A on August 4th 1984.

Now, two satellites are in geostationary orbit respectively at 8° and 5° west, one of the military payloads being a back up for the other. The primary role of Syracuse is to provide communication links between deployed warships and naval ashore authorities, as well as between national authorities and both French forces deployed outside the French territory and military authorities in French overseas territories.

In addition, Syracuse serves as a redundant space based system to the national infrastructure network and battlefield communication circuits.

The ground segment is organized around central stations located in mainland France; these centres are fixed and connected to the Defence terrestrial network. The other stations, scattered over the satellite coverage, communicate with these large mainland stations.

Secure telex and data trafic, at 75 bits/sec, 2.4 and 16 kbits/sec, is handled via modems at each terminal. Telephone links use either a 2.4 kbits/sec enciphered vocoder or 16 kbits/sec delta voice circuits. Spread spectrum techniques are used to provide high resistance to jamming and to enable multiple access to the network.

The TT and C subsytem operates in two frequency bands: S band is used during transfer orbit and also as an emergency back up when the satellite is on station; C band is used throughout the life of the satellite after it has been placed on station.

The ground segment of Syracuse includes fixed stations (M), tactical (TL) and transportable (T) ground terminals and shipboard terminals (N).

The main characteristics of these terminals are given in table 9.5.1.

9.5.2 Second Generation - Syracuse II

The Syracuse II system, currently under development is being designed to replace Syracuse I while extending its capacity and improving its protection against enemy action.

As Syracuse I, Syracuse II shall use civilian multipurpose satellites. As Telecom I, the Télecom II satellites shall operate in C band an Ku band for civilian use and in X band for military use.

Two satellites shall be in orbit, respectively at 8° and 5° west and a third satellite shall be a ground spare. The two operational satellites are scheduled to be launched respectively by mid 1991 and beginning of 1992.

The design life of each satellite is 10 years. The accuracy of station keeping is. 0.05° both in north/south and east/west directions.

TT and C operates in S band during transfer orbit and also as emergency back up during normal operation. C band and X band are used after the satellite has been placed on station.

C band commands are authenticated. X band commands are encrypted.

Autonomy of the satellite in the absence of commands in 24 hours. The military payload of Telecom 2 (SYRACUSE II) is made of:

- 5 transparent X band communication transponders
- one X band command and telemetry unit
- one beacon which transmits the clock signals of the satellite and broadcasts operational messages multiplexed with the command and telemetry signals.

The antenna subsystem provides three coverages

- Central Europe (including France mainland)
- Earth coverage (EC)
- Steerable spot beam.

Circular polarization is used (left for the down link, right for the uplink).

The transmit frequency band is 7.250-7.745 GHz for the down link. The receive frequency band is 7.900-8.395 GHz. The main characteristics of the payload are given in table 9.5.2.

The ground segment of Syracuse II will include an increased number of the stations of the existing types of Syracuse I and new types of stations which are:

- fixed stations in France with 18 meter antennas (M).
- light ship-borne stations (NL).
- very light stations installed on Jeeps (VL),
- airborne stations (AE),
- submarine stations (SM).

The main characteristics of the Syracuse II stations are given in table 9.5.3.

The functions of the various Syracuse stations can be divided into two categories.

1) Fixed stations on the french mainland

- communications with other stations (+ transmission of protected messages)
- supply of relay satellite reconfiguration requests
- payload management and monitoring
- network management
- connection with ground network
- supply of objective information on the status of the network
- TT and C.

2) All other stations called "autonomous" stations

- communications with fixed stations and autonomous stations

- reception of satellite clock signals + protected messages
- connection to users
- possible connection to tactical networks (eg RITA or others)

The network interconnections and the access modes are given in table 9.5.5 (nominal situation) and 9.5.6 (jammed situation) (M refers to fixed stations, AU refers to autonomous stations).

9.5.3 Future Prospect - Syracuse III

The design life of Telecom II is such that Syracuse II should be in operation until the beginning of the next century.

R and D is going on in order to define the main characteristics of what Syracuse III should be.

The aim of the various studies is :

- to improve the existing components of the system,
- to increase its capacity,
- to improve its protection against enemy actions.

For the time being the question on what the architecture of the next system should be remains open. A number of topics are being investigated which should lead towards the definition of this architecture.

Improving the protection against enemy action requires a thorough analysis of the threat (present and future). The following topics are being considered.

- threat against the satellite, i.e. destruction or jamming
- threat against the ground stations i.e destruction, jamming of the fixed station, detection, location and jamming of the mobile or transportable stations.

Studies are made to evaluate the potential interest of using orbit other than geostationary in order to cope with the increasing congestion of this orbit while keeping the requirements for antenna tracking to their minimum and also improving communications at northern latitudes.

Two cases are being considered:

- highly inclined 12 or 24 hours orbits (Sycomore, Molnya, Toundra) for permanent communications,
- low earth orbit for messages.

Such options may involve requirements for interrorbital or intersatellite (on the same orbit) links. Therefore studies are being carried out to define the main characteristics of these links and particularly whether they should use optics or EHF.

A particular effort is made on antennas, in particular on electronic antennas for electronic scanning, electronic pointing, aperture change, nulling againts jammers.

Other means of protecting the links are being studied:

- use of EHF to improve directivity and avoid frequencies congestion,
- modulation schemes: spread spectrum, frequency hopping,
- on board processing.

All the above studies are expected to converge towards the definition of the specifications of the Syracuse III system in the middle nineties.

9.5.4 Sycomores

The Sycomores concept refers to a land mobile satellite communication system using high inclination 24 hours orbits.

The proposed system is suitable for communications over a relatively restricted geographical area at latitudes higher than 35 degrees. The basic mission scenario considers, as an example, an area extending to the Scandinavian countries, the western limit of Portugal and the western half of Turkey at latitudes above 35 degrees north.

In order to be used by relatively cheap mobile terminals the system is designed in shuch a way that tracking is not required. Moreover, the satellite can be viewed by any mobile located in the area defined above at an elevation angle higher that 55 degrees so that the obstructions due to mountains or to buildings in urban areas which would mask a geostationary satellite in the same geographical area are greatly reduced. Since ground antennas having a gain of the order of 10 dBi can be used and since they are pointing at relatively high elevation angle, interference with other services such as line of sight microwave links or communications with geostationary satellites can be easily avoided.

Permanent service is possible with two satellites located on similar orbits having the following characteristics

Period	: 24 hours
Eccentricity	: 0.35
Inclination	: 60 degrees
Apogee altitude	: 50540 km
Perigee altitude	: 21.030 km

The respective ascending nodes of the satellites are spaced by 180 degrees.

On board the satellite, an automatic antenna pointing system tracks a ground beacon in order to keep the antenna beam pointed towards the center of the coverage area.

G/T of the ground terminals is planned to be within the -10 to -15 dB/°K range and the EIRP within the 10 to 15 dBW range. The proposed modulation is QPSK. The proposed access mode is SCPC for the forward link (mobile to satellite)

For the return link each ground switching center transmits on one frequency and TDMA is used for selecting the mobiles located in the relevant area.

The authors of the proposal claim that the communication capacity of the Sycomores system would be 5 times higher and the cost per channel 4 times lower if compared with a similar systems using geostationary satellites:

9.5.5 Reference

[9.5.1] "Sycomores Concept"
J.F. Dulck, R.Larregola, G.Mariet, D.Houffet
Communications and Navigation - London, 17-19 Oct. 1988.

Table 9.5.1

Type of Station	Antenna diameter	G / T dB. K^{-1}	Transmit power	EIRP dBW	Nb of channels
Fixed	8.0 M	31	2 x 1.5 kW	80	2 x 40 MHz
Transportable	3.0 M	22	1.5 kW	74	1 x 40 MHz
Tactical	1.3 M	13	150 W	57	1 x 40 MHz
Shipboard	2 x 1.5 M	17	1.5 kW	67	1 x 40 MHz

Table 9.5.2

Channels	X_1	X_2	X_3	X_4	X_5
Bandwidth (MHz)	60	40	40	40	80
Transmit power (watts)	20	40	40	20	20
Transmit coverage	Center	EC	EC	Spot	EC
	Europe			beam	
EIRP (dBW)	42.4	30.7	42.2	27.7	
at boundary of zone					
Receive coverage	Center	EC	Center	Spot	EC
	Europe		Europe	beam	
G/T (dB/°K)	+ 2.6	−12.5	+ 2.6	+ 1.8	− 12.5

Table 9.5.3

Station type	Antenna diameter (M)	Antenne gain (dBi) Transmit	Receive	G/T (dB/°K)	EIRP at saturation (dBW)
M	18 8	61.0 54.0	60.3 53.3	38.0 31.0	90 81
T	3	44.5	43.8	21.0	74
N	2 x 1,5	39.5	38.8	15.0	67
TL	1.3	38.5	37.8	15.0	65
VL	0.9	35.1	34.3	8.5	43.5
NL	0.9	34.0	33.2	7.5	50.8
SM	0.4	28.0	27.2	− 4.3	38
AE	0.4	TBD	TBD	TBD	TBD

Table 9.5.4
Stations distribution in Syracuse II

	M	N	T	TL	NL	SM	AE	Total
Existing Syracuse I stations	3	11	9	3	0	0	0	26
Added Syracuse I stations		6	8	17				31
New types of stations	(3*)				10	6	26	42 +3*
TOTAL	3	17	17	20	10	6	26	99

(*) Equipment added to existing stations

Table 9.5.5

Nominal

Coverage	Channels	Access Mode
Center Europe	X_1 A11 < - - - > A11	FDMA
EC	X_2 M < - - - > AU	SSMA
EC	X_3 M - - - > AU	FDMA
Spot beam	X_4 AU < - - - > AU	FDMA
EC	X_5 M < - - - AU	FDMA

Table 9.5.6

Under Jamming

Coverage	Channels	Access Mode
Center Europe	X_1 A11 < - - - > A11	FDMA
EC	X_2 M - - - > AU	SSMA
EC	X_3 M - - - > AU	SSMA
Spot beam < - > EC	X_4 AU - - - > M	SSMA
EC	X_5 AU - - - > M	SSMA

9.6 SYSTEMS OF THE FEDERAL REPUBLIC OF GERMANY

9.6.1 Introduction

The Federal Republic of Germany (FRG) has deployed over the last decades a powerful terrestrial communication network based on microwave line-of sight, cables and optical fiber links. This infrastructure can satify the communications requirements of the republic including Berlin (not recognizing the recent merger with the German Democratic Republic). This situation never raised the need to establish a national communication satellite system. The existing communication satellites TV-SAT (18/12 GHz broadcast service, BSS) and DFS-KOPERNIKUS (14/12 GHz fixed satellite service, FSS) have been deployed, more or less, after an initiative of the German industry which had in mind the promotion of the national technology in this matter. This is why the current plans for satellite communications through a national network are pretty weak relative to those of some other European countries.

The following briefly reflects the emerging pertinent ideas in the FRG. The concepts are not necessarily governmental.

At present there is no concept of an advanced DFS-KOPERNIKUS. Two future concepts are in the definition phase: KEPLER1 defined by the German Telecom, and LOOPUS defined by Messerschmit-Bolkow-Blohm.

9.6.2 KEPLER 1

The concept for KEPLER1 anticipates the deployment of a composite payload for BSS and FSS in the SHF frequency range. The key performance features are:

- EIRP	56 dBW (BSS) and 59 dBW (FSS)
- Position	28.5 degrees, geostationary
- Capacity	14 FSS channels 20 BSS channels
- Coverage	1. German language area 2. Central Europe
- Frequency registration	22th Dec. 1988
- Start of operation	1966

9.6.2 LOOPUS

LOOPUS is a communication system concept developed by the German industry and initiated by Messerschmitt-Bolkow-Blohm (Deutsche Aerospace GmbH) in the mid 80 s. It is to establish a Public Mobile Satellite Network (PMSN) with a coverage area in Europe, North America and Asia. The integrated services will be compatible with GSM (Groupe Speciale pour les Mobile), i.e. voice, data etc. with full roaming capability in the service areas.

9 satellites are required to supply the areas from highly inclined elliptical orbits with an elevation angle of greater than 65 degrees for the low cost user terminals achieving almost direct line-of-sight communications.

The frequency range will be strictly 14/11 GHz and 14/12 GHz for the feeder link and the mobile links, respectively, thus re-using the band allocated for FSS from geostationary orbits.

2000 full rate channels at 32 kbps shall be provided by each satellite. The system is reconfigurable to 2000, 4000 or 6000 channels per region.

The interconnectivity features direct satellite switched interconnectivity between subscribers and connectivity to all national public networks in the service areas via up to 40 dedicated Mobile Switching Centers. The current concepts anticipates a TDMA through flying spot (spot diameter 1000 km at present); initial access will be via special global access channel (ALOHA).

Regenerating transponders with on-board processing will provide for circuit switching, antenna control, TDMA timing control, and real-time orbit calculations.

The system is aiming to be deployed in the late 90 s.

A fuller description of LOOPUS is given in Appendix 9 B.

9.7 ESA SYSTEMS

9.7.1 Introduction

The European Space Agency, ESA, has not taken the role of commercially operating a communication satellite system. The established task of ESA and their technical branch ESTEC is to initiate and harmonize future technologies in Europe. ESA is assisting the European industries in the development of advanced technologies needed on the competitive world market of communications satellites.

Future and advanced technologies also encompass the evolution of communication systems and networks. Therefore ESA is establishing so called pilot projects where the industrial community can use their developed equipment on dedicated satellites. Representative derivatives in terms of follow-on commercial programmes have been the INMARSAT satellites and the EUTELSAT spacecraft (the same applies for METEOSAT in the meteorological field).

The current European satellite technology in orbit has now reached an attractive grade of maturity. The pertinent satellites incorporate transparent transponders. Technological upgrades are continuing such as:

- on thrusters
- linearized solid-state power amplifiers
- low noise front end amplifiers
- extension to 20 - 30 GHz
- miniaturisation of microwave and digital technology (MMIC, ASIC)

ESA realised that satellite communications can compete the innovative terrestrial digital communication only if satellite technology makes a similar quantum step as in the terrestrial world. This can be accomplished only by applying novel communication concepts and architectures. For the next generation of communication satellites, ESA is aiming at technologies for regenerative transponders with on-board routing capability ('switchboard in the sky'), mobile communications and optical communications in ISL and IOL. For this objective, ESA has defined a dedicated telecommunication programme which, in turn, is partitioned into three branching programmes.

A) The Payload and Spacecraft Development and Experimentation (PSDE) Programme

The PSDE programme foresees a couple of satellites to be launched before 2000. (There are no plans elaborated yet which go beyond this year.)

The PSDE programme has two main objectives:

- a launch in the mid 90's of a Technology Mission (TM, formerly SAT-2 and now named ARTEMIS) the payload of which has recently been identified but is still subject of modifications due to financial problems.

- a second series of launches in the second half of the 90's, designated as ARCHIMEDES, is a mission on a highly inclined and elliptical orbit.

- a third series of launches in the late 90's, presently *referred to as AOTS-1 and -2 (advanced orbital test* satellite). The planning is likely to be significantly delayed due to financial problems imposed by the budget consuming large programmes as COLUMBUS, HERMES and ARIANE.

B) The Advanced Systems and Technology Programme (ASTP) which contributes to the technology support needed by the PSDE programme. This programme will not be detailed here because it does not encompass a comprehensive communication system.

C) A first operational Data Relay Satellite (DRS) System as part of the future in-orbit infrastructure scheduled for 1996. A Data Relay Preparatory Programme (DRPP) has been initiated to establish and define DRS space and ground segments, investigate feasibility/benefits of interoperability with other DRS systems such as the American TDRSS, and initiate required techonology developments.

The payloads considered for inclusion in the PSDF missions are tabulated in Table 9.7.1 below:

1. Aeronautical/maritime payload
2. Land mobile experimental paylaod
3. Optical IOL payload for GEO/LEO orbit
4. Optical ISL payload for GEO/GEO orbit
5. Single access S-band payload with large antenna
6. Multiple Access S-band payload
7. Propagation payload at 40/50 GHz and/or 80/90 GHz
8. Onboard processing payload

Table 9.7.1 Candidate payloads for the Technology Mission

The contents of supporting technology programmes has been streamlined with the needs of PSDE. Those technology programmes are TRP (Technology research programme) and ASTP (Advanced space technology programme). ASTP has passed 2 phases each about three years long. The current ASTP-3 expires in 1991; ASTP is being planned now to run another three years after mid 1991. To some extent, the TM-satellite will serve as technology demonstration of relevant payload items foreseen for the DRS (Data relay satellite system).

The payload of AOTS is not yet defined. It is very likely that the payload is focussing on regenerative and processing architectures.

A fuller account of the ESA Telecommunication Program Planning is given in Appendix 9.C.

9.7.2 ARTEMIS (within PSDE)

It will be recalled that the Technology Mission ARTEMIS will be based on:

1. SILEX; optical data relay and intersatellite link experiment
2. SDR, S-band high gain, multiple access data relay payload
3. LLM, advanced L-band land mobile payload (successor for the cancelled ARAMIS programme)
4. Spacecraft technology experiments such as ion propulsion,
5. OBP, on-board processing system

where OBP is optional and depends on funding and available spacecraft resources. Additionally, a propagation experiment at 45, 90 and 135 GHz is included. At present, the employment of the existing platforms EUROSTAR or ITALSAT is under consideration.

Spacecraft platform and payloads have a design lifetime generally in excess of 10 years. The experimental phase is to take about 3 years. After that the data relay related capacity of *the Technology Mission may, if required, be used to contribute to* the operational DRS network.

The orbit location will be geostationary over central Europe. Spacecraft operations will be conducted from the ESA station in REDU (Belgium), which will also be the coordinating center for in Orbit Test and Experimentation (IOTE) activities.

SILEX

SELIX will provide a 65 Mbps data capability between an optical terminal on the French spacecraft SPOT 4 and a similar terminal on the Technology Mission. A feeder link in RF will be established between the GEO spacecraft and the SPOT image processing center in Toulouse so as to provide a useful pre-operational service in favour of CNES. An optical ground station is also under investigation for check-out of the GEO terminal and communication experiments.

Another effective application of the optical data relay links is the ISL. The optical terminal on the Technology Mission allows for such an experiment, and the possibility of cooperation with other agencies (NASA; NASDA and INTELSAT) is persued. The feasibility of conducting ISL experiments using an optical ground station is also being analysed. These experiments would aim at verifying the performance of system and equipment such as an on-board processor installed in the ground station.

More details in defining those experiments are going to be elaborated in the coming months.

SDR

The SDR is an experimental forerunner of the S-band payloads which will be required for the future generation data relay systems. It will be used to demonstrate a number of services that may be supported by the future systems before the operation of the Data Relay System.

The SDR payload will be capable of providing a single forward *link from a user earth terminal up to 300 kbps. In the return link,* the SDR can support two users of up to 3 Mbps simultaneously, each tracked independantly over the field of view. For low data rate users with omni-directional antennas, IPFD limitations require spread spectrum techniques and these users will be assigned the same channel. A high data rate user using a directional antenna, will probably be assigned a dedicated channel to avoid potential interference.

The SDR payload will also be capable of providing range and range rate measurements of the LEO satellite. Tracking of the LEO satellite, however, will be performed in an open loop mode using the ephimeries of the LEO satellite to generate steering instructions to programme the phase adjustment of the SDR.

A number of potential users have been identified for a demonstration. One such is SPOT 4.

A user space terminal has been identified based upon spread spectrum transponder developed under ASTP, and an existing omni-directional antenna design. Such an experiment would be valuable in connection with SILEX, offering continuous and *enhanced telemetry coverage and ranging data.*

Also NASA is about to use any of the suitable LEO missions currently supported by TDRS. Other ESA missions under consideration are EURECA re-flights and ARISTOTELES.

The SDR is based upon phased array technology which has not yet been developed and space qualified in EUROPE. In order to provide the high gain required, the large receive array will be deployable in form of two or more panels requiring an innovative mechanical and electrical design. Radio frequency interference

and effectiveness of beam nulling are also areas where ESA discovered the need for in-orbit experimentation.

The test and experimentation programme will address payload characteristics such as communication link performance measurements, beam control, power flux density measurements, RF interference mapping and beam nulling capability, inter-payload interference, and ranging over the user spacecraft. Novel and difficult interfaces such as the communication and coordination between control centers would also be exercised. These aspects include link acquisition, mission coordination, frequency control and timing synchronisation, and user orbit determination.

LLM

The LLM payload will provide for the development and space qualification of much of the technological needed by the high capacity missions for future services. The antenna, presently based upon the Inflatable Space Rigidised Structure (ISRS) technology developed under ASTP, is probably the most critical item requiring in-orbit verification. Of particular interest are considered by ESA the deployment, structural stiffness and dynamic behaviour, reflector accuracy and PIMP's (passive intermodulation products) as a function of ageing effects.

The LLM payload is somewhat related to the planned EMS (European Mobile System) payload that EUTELSAT has foreseen for on their future satellites. (Presently EUTELSAT has discarded the EMS idea, which does not mean that this decision holds for ever). The LLM payload shall be capable of providing spare in-orbit service capacity for EMS. The payload also shall demonstrate the performance of various frequency re-use schemes, including the use of polarisation re-use. Further, the LLM shall demanstrate traffic routing flexibility to the various spot beams. Finally, the LLM payload shall be compatible with various types of services, in particular,

- low data rate service (PRODAT, standard C)

- Private Mobile Radio Service (PMR) based on direct access to the satellite from the customer premisses through VSAT hub stations, for voice and data services

- *Public Telephony Service.*

For the PMR service, studies are under way to define the system architectures compatible with a smooth transition to spot beam systems. Test and demonstrations are needed particularly for an advanced access technique currently being investigated, the synchronized CDMA system, offering an increase in spectrum efficiency. The use of spot beams implies additional requirements on payload performance capability (e.g. differential delay stability between beams) as well as different system monitoring techniques.

The public telephone land mobile service will require an improved quality with respect to PMR services, specifically a higher EIRP per channel, above the level foreseen for the EMS. One approach is a system which is compatible with the Pan-European "GSM" cellular system. Preliminary studies performed by ESCO and ABB in Europe have demonstrated that system integration might be feasible with a payload like LLM, where system integration means that modulation and access techniques are the same in both, the satellite system and in GSM.

OBP

Currently satellite communication using small earth terminals is transmitted almost exclusively through a central hub station. This will not be sufficient for the future integrated terrestrial-satellite communications, which offer a much wider range of applications. For these services on-board processing and switching techniques will be required. In turn, this implies changes in the transmission and access techniques. The acceptance of such new concepts by potential operators and users is dependent on the availability of pertinent technologies and demonstrations.

Before placing a processor in orbit, ESA established in the PSDE programme the so called "Double Hop" experiment so as to promote the liason and cooperation between ESA and the PTT's and EUTELSAT. This experiment is to use an existing transparent transponder from EUTELSAT and to test a processor on the ground along with its control facilities and software.

A simplified but representative OBP system, proposed for flight with the Technology Mission, consists of an OBP payload, a ground based master control facility, and several traffic stations. Particular emphasis has been given to digital baseband technology, system monitoring and control routines, and low cost TDMA equipment.

The introduction of the first operational OBP system in Europe is likely to coincide with the next generation of international communication satellites foreseen by EUTELSAT and INTELSAT. These new systems are about to initiate system definition studies with decisions on system design to be taken during the 1993 to 1995 time frame and launch of the first satellites in the late 1990's.

Other interesting areas of cooperation with EUTELSAT and INMARSAT, such as promotion of ISL aplications in connection with OBP, will be pursued.

Technology Experiments

Briefly the areas of interest for each technology experiment are as follows:

- Ion Propulsion is to reduce the station keeping propellant *requirements of geostationary satellites by some 90%*

- Nickel-Hydrogen batteries from European manufacturers to offer mass saving and increased eclipse capability and to obtain space qualification through the Technology Mission.

- Star Sensor used for, among others, SILEX to minimize its acquisition time.

- Diagnostic Package to provide information on the environment, specifically, RF levels micro-gravity vibration levels, electro-static discharge, and erosion and pollution of optical surfaces.

- Improved Propellant Gauging techniques for more accurate propellant management and mass saving.

- Correlation and assessment of the consistency of Orbit Determination of the Technology Mission and user satellites using several different techniques, e.g. SILEX, SDR, and ESA's new spread spectrum ranging system in preparation for the operational DRS.

9.7.3 ARCHIMEDES

ARCHIMEDES is the designation of a communication system using satellites on a highly inclined and elliptical orbit. This mission is in an early state of definition. ESA studied two inclined orbits (Molniya, Tundra). Follow-on studies are under way so as to select the most suitable orbit in view of the ESA requirements. Initial studies anticipated land mobile and navigation services. Different mission scenarios have been studied ranging from 100 to 5000 channels of 9.6 kbps. The preferred mission is based on a payload with a transparent transponder for 100 channels, 3 spot beams (2.5 m antenna) and 3 satellites on a Tundra orbit. The power consumption of the payload is around 400W, and the

launch mass is about 1200 kg.

Current studies are now focusing on digital broadcast missions. Combined coding/modulation schemes are under consideration as well as the spectrally efficient COFDM (Coded Orthogonal FDM). Operational frequency shall be within 1 to 2 GHz.

9.7.4 AOTS

The AOTS missions are not defined yet. Roughly one can foresee AOTS-1 with a payload providing on-board signal processing (routing) for fixed and mobile services, reconfigurable payloads, and high capacity intersatellite links in optical and millimeter waves. AOTS-2 is likely to consist of a large integrated services communication platform in geostationary orbit, using techniques possibly developed for COLUMBUS and carrying payloads with very large antenna systems, more advanced optical intersatellite links and on-board signal processing for integrated services.

Techniques such as electrical propulsion for orbit maintenance and possibly the use of high critical temperature superconductors in payloads would be experimented in orbit, thus leading to still more efficient and economical communication systems.

9.8 USSR SYSTEMS

A brief discussion of USSR satellite systems is appropriate if only because this nation launches far more satellites per year than any other (74 successful launches in 1989, up to 100 in earlier years). Close monitoring of Soviet space activities is important for NATO not only because of the threat and inconvenience which some of their systems represent but also because of the possibility that a re-shaped NATO could eventually exploit the Soviet capability to launch both large and small payloads into a variety of orbits reliably and at short notice. (Performance of Soviet launch vehicles is tabulated in Section 8.13.)

The USSR has satellite systems for a wide variety of missions including communications, broadcasting, navigation, SIGINT, photo-reconnaissance, missile launch warning and meteorology. Many of these are restricted to low earth orbits where the satellites have lifetimes of only a few months, but extensive use is made of geostationary and inclined highly-elliptical orbits as well. About 70% of the satillites are believed to have a military mission.

Soviet communications satellite systems include RADUGA and GORIZONT in geostationary orbit and MOLNIYA in inclined highly-elliptical orbits. These systems use the civil SATCOM frequency allocations at 6/4 GHz and 14/11 GHz but are known to carry military as well as civilian traffic.

Some Soviet satellite systems give rise to serious potential interference problems for NATO and the allies. Although the systems mentioned above do not use the 7/8 GHz frequency band relied upon by the allies for military SATCOM, the USSR has since the mid-1970s announced its intention to deploy a geostationary MILSATCOM system called GALS which would operate in this band. To date some twenty GALS locations around the geostationary arc have been declared to the IFRB and formal coordination with NATO nations proposing satellites neaby has been required. As far as is known, however, no GALS payload has yet been put into orbit.

More recently, the Soviet Union has announced a new SATCOM system called TDR, which would operate in the EHF band with 44 GHz uplinks and 20 GHz downlinks. Unlike allied EHF systems such as MILSTAR which have on-board processing, TDR would provide a number of transparent channels linking mobile terminals to fixed anchor stations. In this case no less than 25 geostationary locations have been declared to the IFRB, but again, no TDR payload has yet been orbitted.

We may postulate that in declaring systems like GALS and TDR to the IFRB the USSR is simply earmarking a large number of positions on the geostationary arc to ensure that should a need arise to deploy an SHF or EHF COMSAT in the future it could do so without hinderance and with minimum constraint on location. It is fairly straightforward to change the parameters in the system declaration once a "slot" has been secured. It is even possible that the USSR has no intention of using either the 7/8 GHz or the 44/20 GHz frequency bands but is declaring "paper" satellites purely to create coordination difficulties for the allies, and perhaps force them to non-optimum locations or to concede EIRP and/or bandwidth reductions.

Some Soviet meteorological satellites also use frequency bands that overlap those used for allied SATCOM. A particular case is GOMS-1M which, in the form originally declared to the IFRB, would deny NATO the use of 50 MHz of bandwidth on NATO IV at 18 degrees W. Negatoations are currently in progress to try and resolve this problem.

The Soviets have a UHF SATCOM system called VOLNA which supports mobile terminals, in particular shipborne terminals. The mobiles communicate via fixed anchor stations whose links to the satellites are at L band. VOLNA UHF frequencies are contiguous with those used by allied UHF MILSATCOM systems such as FLTSATCOM and those of NATO IV and Skynet 4 satellites. Although specific frequency allocations have been made to individual systems there is clearly scope for mutual interference. Unlike GALS and TOR, VOLNA is an operational system. The VOLNA space segment does not consist of independent satellites but takes the form of a set of additional payloads on SHF geostationary satellites such as GORIZONT

The Soviets also use small, low-orbit satellites for communications and have for many years operated a UHF store-and-forward system of this kind for message traffic. The satellites are launched in groups of eight.

APPENDIX 9A

THE FIRST "SWITCHBOARD IN THE SKY" : AN AUTONOMOUS SATELLITE-BASED ACCESS/RESOURCE CONTROLLER*

Satellite communications systems can connect hundreds of dispersed users instantaneously. But when such a system is serving many small, widely dispersed users and is handling fluctuating traffic loads, the satellite channels must be allocated dynamically. A satellite - based access/resource controller, a "switchboard in the sky, "has long been seen as the most efficient way to use the costly resources of a satellite. Two satellite communications packages, which will provide highly protected links to small mobile terminals, incorporate the first autonomous satellite - based switchboard.

Satellite communications offer a compelling advantage for military and civilian applications-instantaneous connectivity among hundreds of widely dispersed and mobile users. Military users, for example, use satellite communications to link ships, aircraft, submarines, even individual soldiers. Civilians communicate around the world and with remote villages via satellite.

But current satellite communications technologies have several serious deficiencies. Chief among these, for the military, is an extreme vulnerability to deliberate interference, "jamming." And both military and civilian users need a solution to the Demand Assigned Multiple Access (DAMA) problem.

When a satellite communications system is serving many small, widely dispersed users with widely fluctuating traffic loads, satellite channels must be allocated in a highly dynamic way (to utilize most efficiently the very expensive communications capacity of the satellite). This is the DAMA problem. The DAMA service available for today's satellite communications technologies is not totally satisfactory. A satellite-based access/resource controller, or"switch-board in the sky," has long been seen as the ideal solution to the DAMA problem.

In a new type of satellite communications package, called FEP (FLTSAT EHF Package), we have addressed the problems of jamming, DAMA, and other military-specific problems of satellite communications. The FEPs ride on two FLTSAT satellites, FLTSAT-7 and FLTSAT-8. These communications packages can provide highly protected satellite communications to many small mobile terminals at extremely high frequency (EHF).

The FEP features a true "switchboard in the sky" function. To our knowledge, it is the first of its kind.

CURRENT SATELLITE SYSTEMS

A typical communications satellite uses transponders, which simply receive uplink signals, amplify them, and retransmit them at a different frequency on the downlink. The use of different uplink and downlink frequencies prevents the satellite's transmitter from jamming its incoming signals. By using multiple access techniques, multiple terminals can share the communications resources of each of a satellite's transponders.

Two common types of multiple access are frequency-division multiple access (FDMA) and time-division multiple access (TDMA). In an FDMA system, the satellite's receiving frequency band is divided into narrower sub-bands (channels). Each user's terminal transmits data over a carrier frequency within its assigned channel. Data are transmitted by modulating the carrier frequency. In this way, the frequency "carries" the data. Since FDMA systems have multiple channels, they use multiple carrier frequencies

A TDMA system uses only one carrier frequency. Access to the system is based on a time period or data frame, which is divided into time-slots. Terminals are assigned particular time-slots, during which they can transmit and/or receive data. The TDMA scheme requires users to be synchronized in time-so they know when their opportunities to transmit and receive data occur.

The FDMA frequency channels and TDMA timeslot channels make up a communications resource poll that must be allocated to user communications links. The allocation process, therefore, must include a mechanism for informing each terminal of its channel assignment.

The communications resources of satellites can be assigned statically or dynamically. Statically assigned satellites either have fixed channelization, which locks in spaceceraft resources without regard to traffic conditions or are reconfigurable via commands from a central ground-contral station (Fig. 1). Reconfiguration from a ground-control station is relatively slow, but this method provides adequate service when traffic is fairly constant.

When communications traffic fluctuates widely, channel assignments must be made dynamically-a DAMA system is required. A DAMA system allocates and releases communications channels in response to varying user traffic demands. When the level of communications traffic is intermittent, a DAMA system can support many more users that a statically assigned system.

All DAMA systems have a control channel, which terminals use to request communications service and receive channel assignments. In a ground-based DAMA system using a transponding satellite, one terminal can act as a central controller or a distributed control protocol can driect terminals to available channels. The SPADE (Single Channel Per Access Demand Equipment) system currently in use on INTELSATs (International Telecommunications Satellites) uses a distributed control protocol. Central control is appropriate when allocating channels of different capacities. Such systems serve civilian needs quite well today, as long as all user terminals are in the antenna beam of the same satellite.

But if users are widely dispersed (as mobile users often are), multiple-beam systems make good economic sense. The most precious resource on a satellite is transmitter power. By using multiple narrow beams, transmitter power can be better concentrated on user terminals. As the traffic varies, however, a satellite's capacity must be allocated among multiple beams. In the transponding satellites in use today, entire transponders must switch among beams. But switching an entire transponder as traffic varies is inefficient, because it switches a large block (many channels) of a satellite's capacity.

In contrast with the transponding satellite, the ultimate DAMA satellite is a "switchboard in the sky": a satellite that can quickly allocate individual circuits among its multiple beams in response to direct user requests (in the same way that a central telephone switchboard responds to user requests). In a satellite switchboard system, user terminals send control messages directly to the satellite (Fig. 1). The control messages may request communications service (much like dialing a telephone number). When a user terminal dials a link (by sending a control message), the switchboard allocates spacecraft resources; when

(*)This Appendix was written by D.Semprucci of the Lincoln Laboratory, MIT,USA.

the user terminals hang up (by sending another control message), the resources are released and made available to other terminals.

Until recently, such a satellite-based switchboard was infeasible. Now, it is not only feasible, we have implemented it in FEP and it has been functioning well on-orbit since December 1986. This switchboard has provided communications service to many small mobile terminals. And these terminals have both widely varying antenna sizes and widely varying communications needs.

FEP CHARACTERISTICS

A signal-processing satellite (unlike a transponding satellite) separates all user channels on the uplink, completely demodulates each channel, and repackages the channels on the downlink, usually in a different modulation format. The primary reason for this procedure is to separate jamming signals from valid user signals. However, this procedure gives the satellite the ability to control and assign communications resources on an individual channel basis. Therefore, an onboard signal-processing capability makes a satellite-based switchboard feasible.

The FEP is a signal-processing satellite. It has two beams; its communications links are configurable upon request. Unlike transponding satellites, which must switch large blocks of channels when switching capacity between beams. FEP can switch individual channels between beams. Data from an uplink channel can be routed on the downlink to the same beam, to the other beam, or to both beams.

Like all communications satellites, FEP's transmitter power is limited. But the amount of downlink resources allocated to a particular communication link can be tailored to fit the needs of individual users. The FEP supports widely different sized terminals and, since smaller terminals consume more downlink resources than larger terminals (because of the small size of their receiving antennas), FEP makes efficient use of its transmitter power.

As shown in Fig.2, the FEP's beams are an earth-coverage beam and a 5° steerable spot beam. The spot beam illuminates an area appproximately 2,000 miles wide. The spot beam can be steered via control messages sent directly from privileged user terminals to FEP's satellite-based switchboard.

Although FEP has only two beams, the principles demonstrated apply equally well to systems with many more beams. Each beam in FEP has a 44-GHz receiver. The uplink is frequency hopped for anti-jamming protection. It uses a combination of FDMA and TDMA techniques. The uplink frequency bands are divided into FDMA channels-26 can be used for communications links.

Each of the 26 uplink FDMA communications channels is divided into three types of TDMA timeslots: C0, C1, and C2 (Fig.3). The C0 timeslots are for user primary communications data. C1 timeslots are for user secondary communications (at low rates). The C2 timeslots are used for control messages that are sent from user terminals to the satellite-based switchboard. The C2 portion of each uplink channel is subdivided into additional timeslots, called "accesses," which let up to eight terminals share the C2 portion of one uplink channel.

The FEP's downlink is TDMA using a single hopped frequency carrier in the 20 GHz band. Data are transmitted over the earth coverage beam or the 5° spot beam on a timeslot-by-timeslot basis. Several timeslots on the downlink are fixed; they are devoted to such overhead functions as synchronization (SYNC) signals, telemetry (TLM) data, and access control (AROW/C3) messages (Fig. 3). The remaining timeslots (80% of the downlink) are available for assignment via the switchboard for user communications links.

Each of FEP's 26 uplink communications channels can support up to 2,400 bps (vocoded voice) communications. Additional uplink channels are dedicated to command and control of the FEP itself. To provide communications service to the maximum number of users, FEP's uplink and downlink processors are configurable in a variety of modes, which are all under the control of FEP's autonomous switchboard.

OPERATION

Figure 4 shows the types of signals that pass between FEP and the user terminals. FEP is continuously transmitting a synchronization signal (1) in fixed timeslots on the downlink. At its turn-on, a user terminal must look for the synchronization signal and align its timing to match the satellite's TDMA cycle timing. This process is called "acquiring the downlink." At this point, the user terminal is able to be a passive listener on communications links. However, if the terminal wants to be an active participant in communications links, it must also align its uplink timing. To do this, it sends uplink probes (2) in a preassigned uplink channel and timeslot reserved for acquisition. The timing acquisition processor on FEP (which is contained in the uplink processor) responds with AROW control messages (3) in fixed downlink timeslots, telling the terminal if it is early or late with its probes. The terminal then adjusts its uplink probes and tries again. Once a terminal's uplink timing is correct, the timing acquisition processor on FEP notifies the switchboard that a terminal is acquiring. The switchboard communicates with the user terminals via C3 control messages (5) sent in fixed timeslots on the downlink. Following acquisition, the switchboard sends a C3 control message to the terminal and assigns the terminal an uplink FDMA frequency channel and C2 TDMA timeslot. Now the terminal is enrolled in (logged on) the satellite.

An enrolled terminal communicates with the switchboard via C2 control messages (4) transmitted in its assigned uplink C2 channel and timeslot. For example, the terminal can send a C2 message up to the switchboard and request a communications link. The switchboard makes the necessary connections in the satellite and then sends the terminal a C3 message that specifies the uplink frequency channel and downlink timeslot assignments. The terminal can then begin communicating (6). These transactions are computer-to-computer; the protocol is transparent to the terminal operator and takes only a few seconds.

CONNECTIVITY AND CONTROL

The block diagram in Fig.5 shows the overall design of the FEP. Data enter the FEP through either the earth-coverage receiver or the 5° spot beam receiver. The data are dehopped, demodulated and passed through the signal-processing package, where the data are rerouted, rehopped, remodulated and retransmitted over the 5° spot beam, the earth-coverage beam, or both beams. The uplink and downlink processors contain tables that determine which channels' data will be processed and how to process the data. The switchboard is a microprocessor that controls these tables via an IEEE-488 parallel data bus. If the IEEE-488 bus malfunctions, a slower backup serial bus can be activated.

The switchboard activates channels in the uplink processor and specifies the data rate and other mode information for processing data from the active channels. Until an uplink channel is activated by the switchboard, the uplink processor throws away data from that channel. Once a channel is activated, its C0 and C1 data pass to the downlink processor, where they are stored while they wait for transmission on the downlink. The switchboard then activates downlink timeslots for data transmission and assigns to the downlink processor a beam and a downlink burst rate for transmitting the data. By devoting more downlink timeslots to repeating data, a communication link can

be made more robust on the downlink. Thus the switchboard connects uplink frequency channels to downlink timeslots. Until the switchboard makes this connection, the signal processor discards data from these uplink channels.

Figure 5 shows an example of how a terminal interacts with the switchboard to place a point-to-point telephone call to another terminal. The other terminal can be in the same beam or in the other beam. Moreover, the caller doesn't need to know which beam the other terminal is in. The FEP's switchboard remembers the beam assignment from a terminal's log-on and forms the necessary connections in the satellite and sends the user terminals their channel assignments.

In this example, Terminal 1 and Terminal 2 are in different beams. If Terminal 1 wants to call Terminal 2, it sends a formatted C2 control message to the switchboard - an Initial Service Request (C2-ISR). This message specifies the data rate and tells the switchboard that Terminal 1 wants to talk to Terminal 2. The switchboard checks its data base to see if Terminal 2 is enrolled in the system. If it isn't enrolled, the switchboard sends a C3 message to Terminal 1, notifying it that Terminal 2 is not logged on. If Terminal 2 is enrolled, the switchboard sends a C3 message to Terminal 2 called aRing-Up message (C3-RU), analogous to a telephone ring.

Terminal 2 responds to the "ring" with a C2 message called a Call Answer (C2-CANS) in which it tells the switchboard if it is accepting the call and how robust a downlink it needs. The switchboard then checks its data base to see if it has enough uplink frequency channels and downlink timeslots to form the link. If the channels and timeslots are available, the switchboard activates the channels in the uplink processor and the timeslots in the downlink processor. It then sends the channel assignments to both terminals via a C3 Service Assignment message (C3-SA).

Terminal 1 can now begin transmitting on the uplink. Data from Terminal 1 are processed by the uplink processor, passed to the downlink processor, and retransmitted to Terminal 2. Unless a reconfiguration is needed, the switchboard has no further involvement with the link. The data just pass between the uplink processor and the downlink processor. If a person at Terminal 2 wants to become the active talker, he selects his "push-to-talk" button and his terminal automatically sends a C2 message to the switchboard that request a reconfiguration. The switchboard that then reverses the link, sends both terminals a C3 message, and enables Terminal 2 to transmit on the uplink.

This example describes a half-duplex terminal-to-terminal call. Other protocols in FEP support full-duplex terminal-to-terminal calls and multiuser networks (which resemble conference calls).

The priority level of a call is set by the C2 request that initiates the call. If the switchboard does not have enough available uplink or downlink resources to set up a service, it will check its data base to see if any existing services have a lower priority. If a lower priority call is found, the switchboard will preempt as many of these services as necessary (starting with the lowest priority), send C3 messages to those services that notify them of the preemption, and then configure the higher priority service. The terminal requesting the higher priority service isn't aware of the preemption, and the transaction takes only a few seconds. Thus communications resources don't have to be reserved in this system to ensure that important calls go through.

These protocols are computer-to-computer transactions. The user is no more aware of the details of the switching protocols than a person who makes a conventional long-distance call.

Besides accepting C2 control requests from user terminals, the switcboard can receive the following commands from the FEP Operation Center (FEPOC):

- Upload the parameters that the switchboard uses to make decisions.
- Activate/deactivate special control signals that disadvantaged terminals (especially submarines) use.
- Set up fixed communications links with specific uplink channels and downlink timeslots (like conventional satellites). This command preempts any services using those resources and takes them out of the assignment pool.
- Turn on/off the output of diagnostic information via telemetry.

FEP IMPLEMENTATION

Each FEP package weighs 245 lb and draws 305 W. The electronic boxes that make up the package are mounted on a hexagonal module. This module is attached to the rear of the FLTSAT spacecraft (Fig. 6). The FEP antennas lookthrough openings in the ultrahigh frequency (UHF) transmit-antenna reflector.

Shown in the photographs in Fig. 7 are two FEP modules. The FEP-8 module (in temperature chamber) is getting ready for temperature stress tests; the FEP-7 module (in front) is undergoing testing with its antenna assembly. Figure 7 also shows a typical FEP digital electronics box. The FEP has a separate box for the uplink processor, the downlink processor, the TRANSEC processor, and the switchboard. The boxes are mounted on the inside of the FEP hexagon. The switchboard box consists of seven circuit boards: a processor board, two input/output boards, and four memory boards.

The FEP resides on a satellite, so power and weight limits are stringent. The heavier a satellite, the harder it is to launch. Moreover, power is limited. FEP receives its power from FLTSAT's solar arrays and uses battery backup during eclipse periods. Weight and power requirements were satisfied-the switchboard weighs only 6 lb and draws 6 W.

Beacuse FEP is in synchronous orbit, it is exposed to radiation and therefore the switchboard microprocessor had to be radiation hard. We ran radiation tests on several microprocessors and chose an I^2L version of the Texas Instruments (TI) 9900. This microprocessor is well established and well supported. It comes with a development system that includes a TI minicomputer (used to develop, compile, and unit-test the source code) and a real-time emulator. The emulator was attached to an off-the-shelf NMOS 9900 board. This configuration allowed real-time testing of the switchboard software while the switchboard hardware was being developed.

The switchboard software was written in Pascal, a programming language that imposes structured programming techniques. The source code was compiled under TI's MPP Pascal compiler; it runs under TI's RX operating system. The MPP Pascal compiler and its associated RX operating system include such extensions to the Pascal language as the ability to run a multi-tasking program in real time and to handle interrupts with Pascal processes. (For speed, some Pascal interrupt handlers call assembly language subroutines.)

The MPP system provided a reverse assembler, which showed the assembly language code that the Pascal compiler generated. All of the switchboard's object code was run through the reverse assembler. This output, coupled with the source code for the RX operating system, allowed us to know the contents of every location of the switchboard's program memory. We also wrote programs that analyzed the data base and provided the same location-by-location information for the switchboard's data base. Thus every location in the switchboard's memory is identifiable and can be monitored.

One major concern was that the TI 9900 microprocessor could only address a maximum of 64K. Therefore, both the program and its data base had to fit within the 64K limit. This was a challenge! Using the reverse assembler, we identified areas where the Pascal compiler was inefficient in converting Pascal code to assembly-language code. We then compensated for these inefficiencies in our Pascal coding. In the end, the program successfully fit into the available memory.

SOFTWARE OVERVIEW

Figure 8 gives an overview of the switchboard software. This software is a real-time multitasking system. Each box in the figure represents a separate task or process, which all operate in parallel. The boxes on the top are the input handlers; the ones on the bottom are the output handlers. When a terminal logs onto FEP, the timing acquisition processor (TAP) notifies the switchboard by passing a message. The TAPIN process handles the message. TAPIN looks in the data base for an available C2 channel. It activates that channel in the uplink processor by sending data over the IEEE-488 bus and sending the terminal a C3 message that specifies its channel assignment.

Any C2 messages sent to the switchboard are handled similarly by the LC2IN process. If, for example, the C2 message is a request for a communication link, process LC2IN checks the data base to see if there are enough uplink channels and downlink timeslots to configure the link. If so, LC2IN activates the channels by sending data to the uplink and downlink processors over the IEEE-488 bus. Then the process sends a C3 message to the terminals that assigns them their channels. If the C2 message is a request to move the 5° spot beam and the message is from a privileged terminal, process LC2IN sends the data to the antenna-pointing interface, which moves the 5°dish. LC2IN then sends a C3 message to the terminal, which tells it that the spot beam has moved.

To avoid deadlock, output processes are generally given higher priority than input processes. The only exception to this rule is the MAIN process, which is the command input handler. If a failure occurs on-orbit, we must be able to issue a command to clear the failure and/or dump portions of the switchboards memory over telemetry for analysis on the ground. No failures are expected, but the diagnostic capability is available and has been thoroughly tested.

DATA BASE

The switchboard's data base takes up 25% of the switchboard's memory. Two thirds of the data base is devoted to the stack, which contains statically allocated data such as uplink channel tables and downlink timeslot tables.

The remaining data base, which is called the heap, consists of packets of dynamically allocated data. The heap contracts and expands in response to user terminal activity. It consists of terminal packets (which contain information for each logged-on terminal), service packets (which contain service parameters for each active communication link), timeout packets awaiting C2 responses, and I/O packets for the switchboard's input and output queues. Figure 9 shows examples of stack and heap elements.

We have divided the heap into a reserved section and a general unreserved section. During initialization, the switchboard software reserves sections of the general heap. Each of these sections, called subheaps, is devoted to one type of data base element. There is a subheap of each of the switchboard's input and output queues, a subheap for terminal packets, a subheap for service packets, and so forth.

The reserved heap serves two functions. First, it prevents fragmentation of the data base. The size of heap packets is variable; terminal packets are not the same size as service packets, and so on. Thus as the heap expands and contracts (in response to user activity), the data base memory can become fragmented so that the total memory is large enough for a packet but the contiguous memory is insufficient. By keeping identically sized packets in one subsection, we prevent the fragmentation.

The second advantage of the reserved heap-preventing the system from becoming unbalanced-is even more important than the first. If, for example, too many terminals log onto FEP, sufficient memory for holding service parameters may not be available. If this situation occurs, communication links cannot be set up. So we pre-reserve enough memory for 50 terminals, 26 service packets, and various sizes of input, output, and time-out queues. Then, if a 51st terminal logs onto the system, the terminal can overflow into the general unreserved heap. We keep the system flexible (and don't lock in memory resources any more than the communications resources), yet ensure a reasonable, balanced system. Figure 10 shows the allocation of the switchboard's memory.

SAFEGUARDS

In a satellite-based switchboard system, the switchboard processor must execute its tasks flowlessly. Therefore, we tested the processor exhaustively. But the system must operate in real time and we can't test every possible scenario. So we built an extensive telemetry reporting feature into the switchboard. Using the reporting feature, an operator in FEPOC can monitor the "health" of FEP and its switchboard.

The switchboard is constantly sending indications, via telemetry, of its current state: whether it's at "idle" or is processing an interrupt; its current interrupt level; the state of the IEEE-488-bus or backup-serial-bus handshaking signals; and whether the program is executing a portion of code that could cause a "hangup" if a hardware failure occurred. This last condition would occur if, for example, the bus output handler was active constantly. This activity might indicate a hardware failure of the IEEE-488 bus. The telemetry reporting enables us to detect such a failure on the ground and activate the backup serial bus.

The switchboard can echo any or all of its input and output data via telemetry. If the FEP is in a testing phase, a command can activate or deactivate this feature. When the echo is activated, FEPOC can monitor all the switchboard's interactions in real time. The interactions can also be stored for delogging at a later time. The echo data are useful not only for checking the switchboard functions, but also for diagnosing user terminal problems.

The switchboard can also be commanded to dump portions of its memory via telemetry. These dumps provide information about the state of FEP and can help us diagnose anomalous conditions of the switchboard.

The entire switchboard program is contained in read-only memory (ROM), but it is executed in random-access memory (RAM), which reduces power consumption and enables us to patch the program. The last 512 bytes of the switchboards memory contain a loader program, resident in ROM. A command can activate the loader program, which then downloads the contents of the switchboard's ROMs to the RAMs. An alternate mode of the loader lets us repair errors in the switchboard software by uploading a patch or an entire new program to the RAMs. This alternate mode can also be used to upload new versions of program, possibly implementing different protocols and thus extending the useful life of FEP.

If the switchboard hardware fails, a command can set the switchboard into a dead state. In this state, special hardware reads switchboard commands and puts the command data directly onto the backup serial bus. Data sent to FEP in this mode can be used to reconfigure the uplink and downlink processors. FEP the operates as a "slowly reconfigurable" conventional

communications satellite under direct control of a ground station.

As another backup option, the uplink and the downlink processors can be commanded into a fixed default configuration. Thus, if there is a fatal failure of the satellite-based switchboard, FEP can still operate as a statically assignable communications satellite. The command channels used to activate these backup modes are protected from switchboard failures-the switchboard cannot reconfigure either of these (one in each beam) dedicated hardwire command channels.

SUMMARY

Meeting the communications needs of many small, geographically dispersed users with widely fluctuating traffic loads has always posed a problem for satellite communications systems. The FEP, with its satellite-based switchboard, provides one solution to this problem. It represents an ideal DAMA system, one that can dynamically allocate satellite communications resources among multiple beams in response to the varying traffic demands of many small, mobile terminals. These terminals have various sizes of antennas and, therefore, require different amounts of uplink and downlink capacity. The FEP can quickly switch communications channels between its beams and can tailor the amount of resources allocated to each user. It, therefore, can make highly efficient use of its communications resources.

The FEP can preempt communications services automatically and prevent important calls from getting busy tones. The switchboard features many safeguards that ensure proper operation.

On 4 December 1986, FLTSAT-7 and its FEP were successfully launched from Cape Canaveral on an Atlas/Centaur booster. After an initial check-out period, FEP began operating successfully as a satellite-based switchboard for the on-orbit testing of user terminals. The user terminals have included small submarine and airborne terminals, as well as larger shipboard and land-based terminals. A small manpack EHF terminal called SCAMP (Single-Channel Advanced MILSTAR Portable) developed for the Army by Lincoln Laboratory has also been successfully tested with the on-orbit FEP. The FEP-8 has been mated with the FLTSAT-8 satellite, which was originally scheduled for launch in May 1987. The launch actually took place in 1989.

The advent of radiation-hard microprocessors and high-level structured programming languages has made it possible to use a communications satellite to handle increasingly complex control functions. The success of FEP has demonstrated that an autonomous satellite-based switchboard is a viable concept and as of 4 December 1986, we have the first "switchboard in the sky."

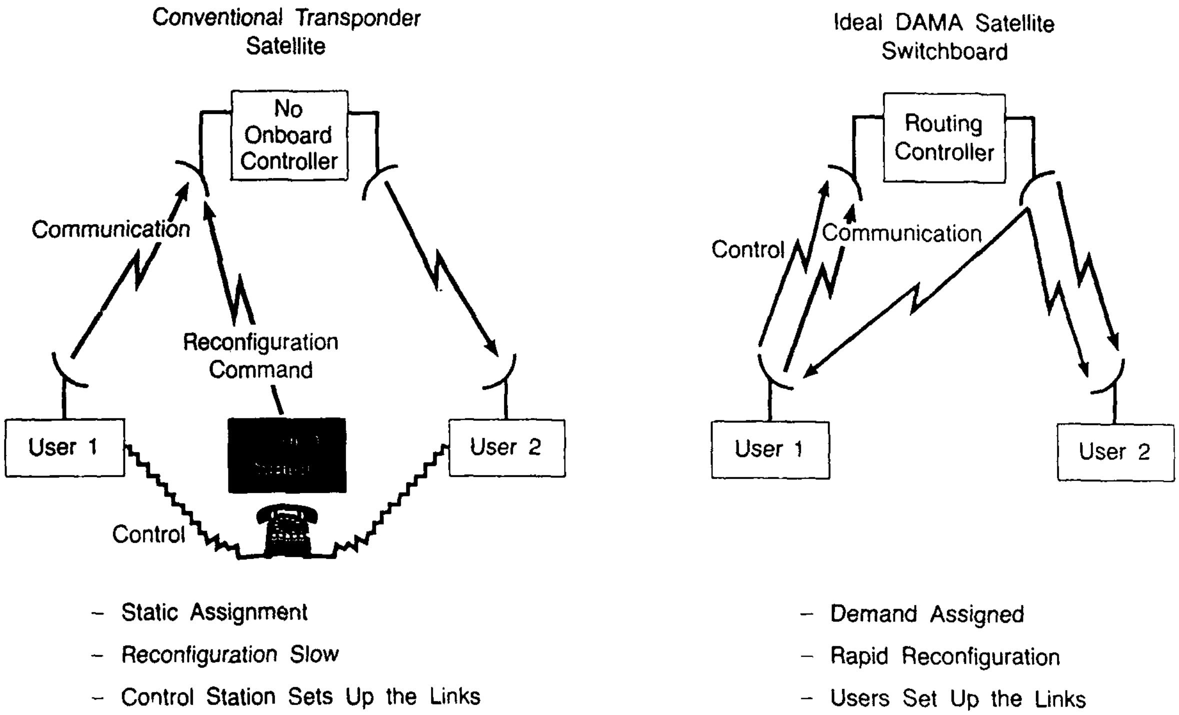


Fig.1 Access control techniques for a satellite communications system

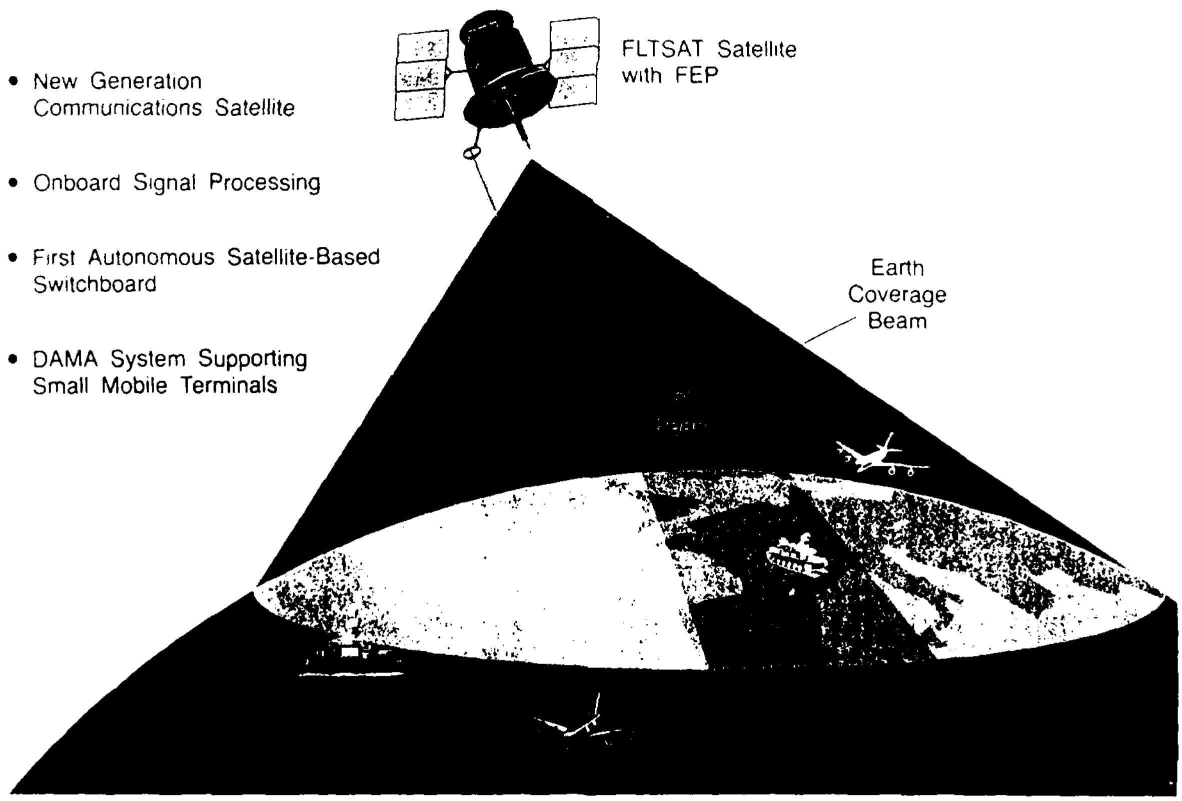


Fig.2 FEP characteristics

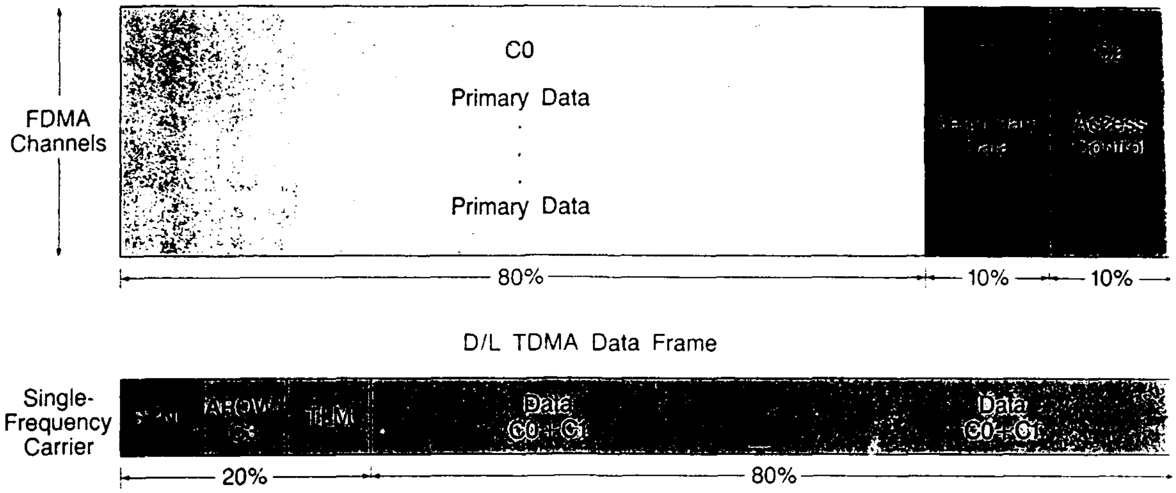


Fig.3 Channel resources (U/L = Uplink; D/L = Downlink)

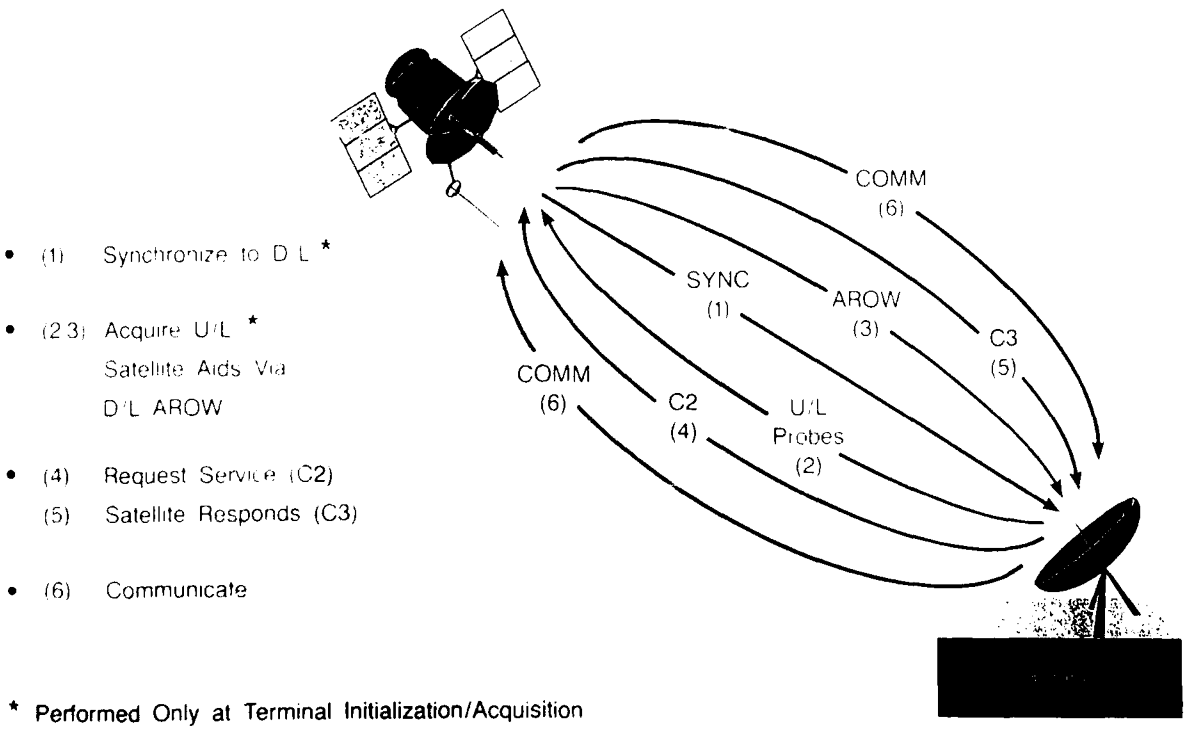


Fig.4 FEP signalling procedures

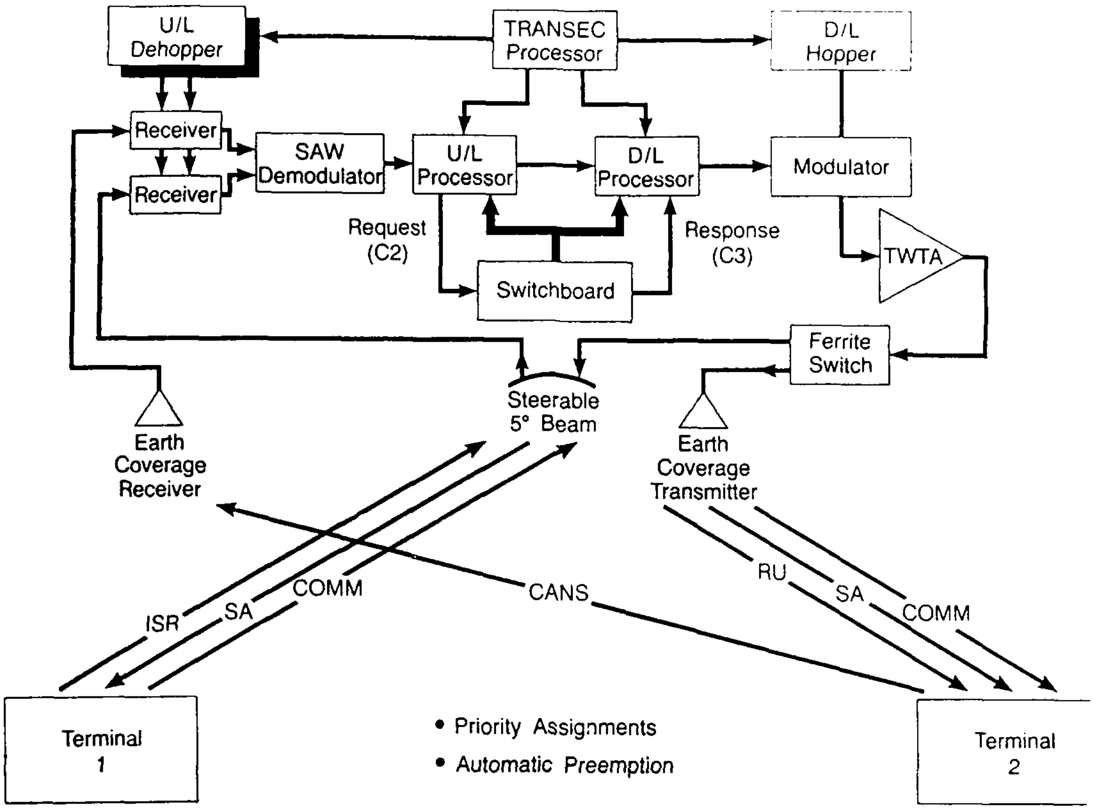


Fig.5 The FEP allocates and releases channels as user traffic varies

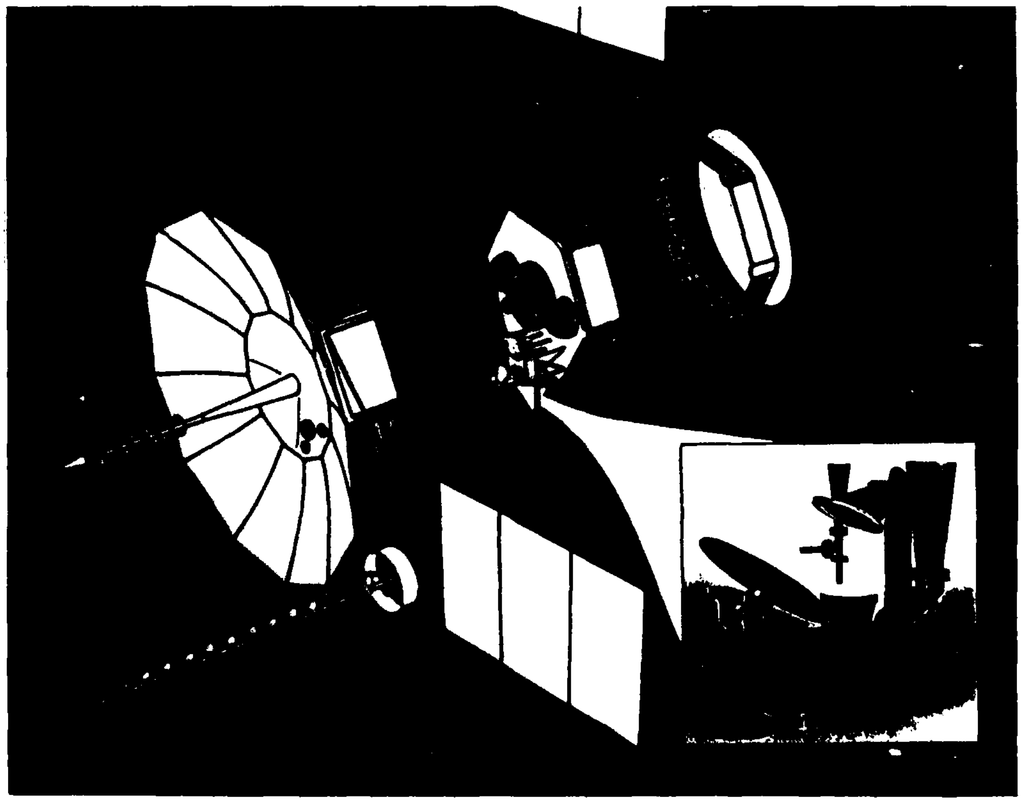
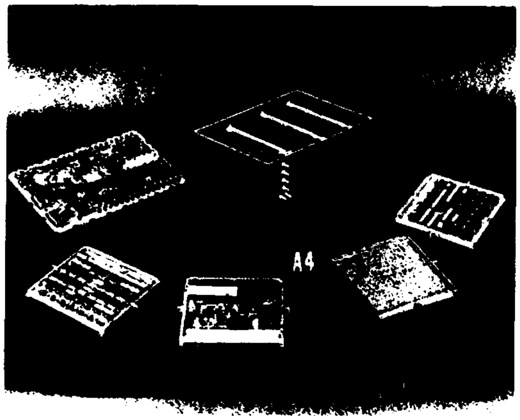


Fig.6 The FLTSAT spacecraft FEP antenna system is shown in the inset



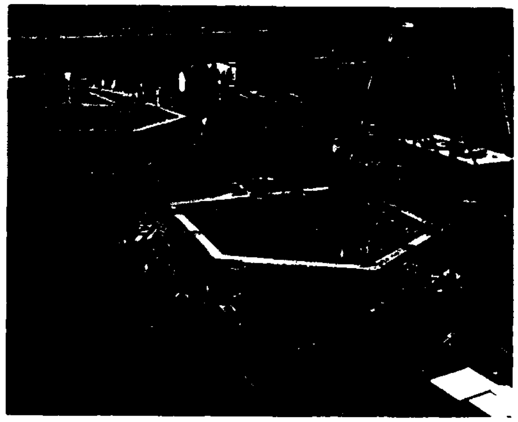


Fig.7 Implementation of FEP. The device includes seven boards, weighs 6 lb and draws 6 W. It includes a real-time emulator, real-time Pascal, and 64K of memory

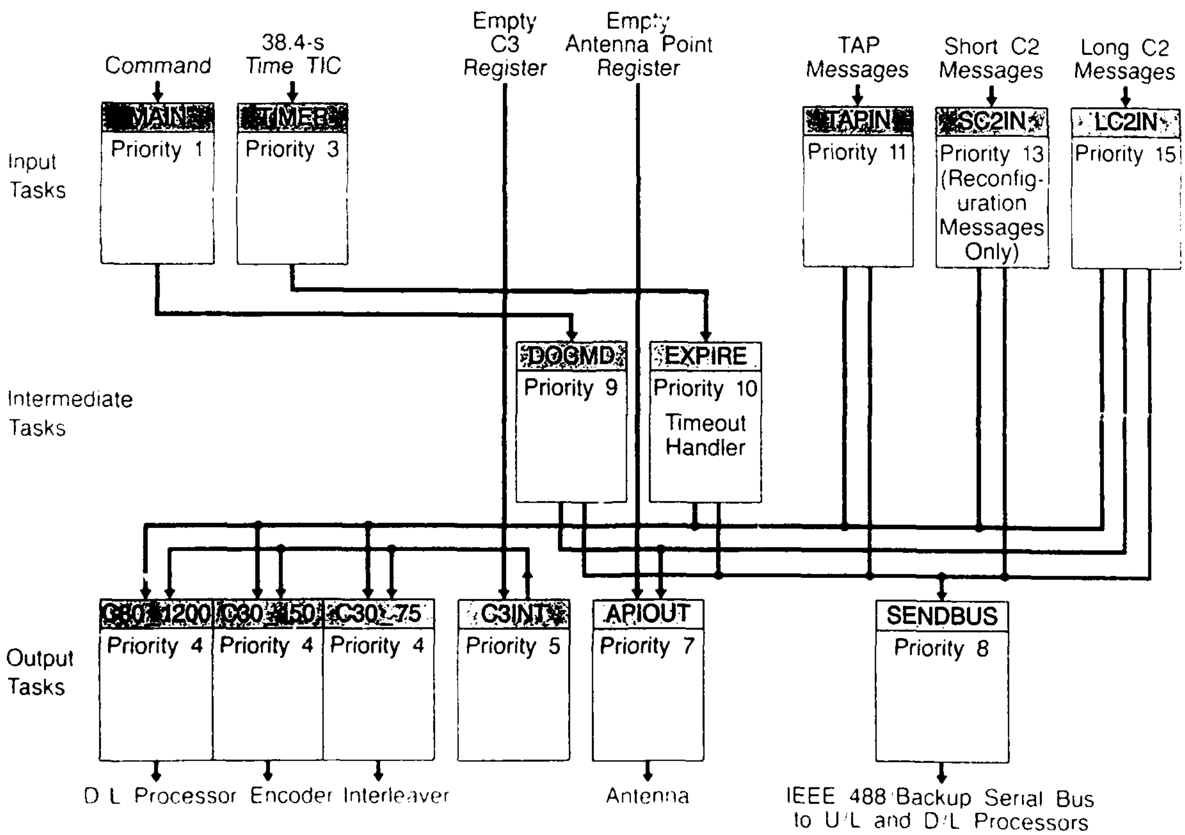


Fig.8 Switchboard software tasks

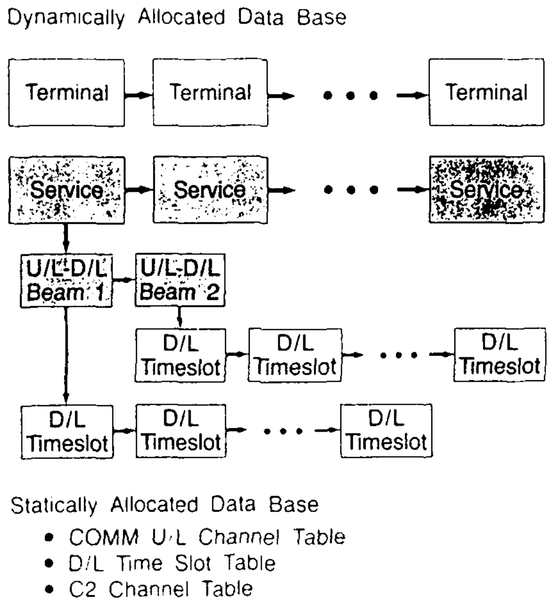


Fig.9 Onboard data base

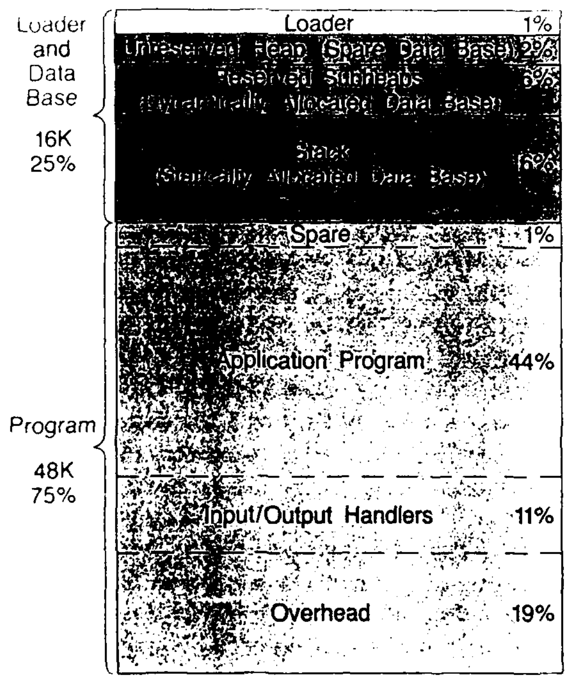


Fig.10 Switchboard memory organization

APPENDIX 9B
LOOPUS

1. Introduction

The geostationary orbit enjoys the considerable merit of the apparent stand-still of its satellite. However, the geostationary orbit also has two substantial constraints. These are the limited number of available orbit positions and the ability to illuminate the polar regions.

Under certain constraints, constellations of orbits inclined againts the equatorial plane, can produce quasigeostationary orbit positions offering properties similar to geostationary satellites, i.e. continuous communication between earth stations round the clock.

A quasi-geostationary satellite system described below is referred to as LOOPUS system. LOOPUS stands for "geostationary Loops in Orbit Occupied Permanently by Unstationary Satellites"..

The initial LOOPUS concept studied under a contract of the Deutsche Bundespost (German PTT, now Telecom) was based on a 12 hours orbit with orbit parameters given in Table 1. While in this concept the fixed satellite service (FSS) played a predominant role, recent follow-on studies have shown that an integrated mobile service (land, air and maritime) would be suited better to the particular LOOPUS system. This alternative concept, designated LOOPUS MOBILE D, adopted a 16-hr orbit thus providing for an improved coverage of land regions.

The frequency band assigned to mobile satellite communication at about 1.5 GHz, is at present very limited. Exploitation of inclined orbits permits the re-use of frequencies using state-of-the art technologies, particularly in the 14/12 GHz band. This solution offers a large bandwidth and good growth potential for future system expansion.

2. Orbit Concept and Earth Coverage

The orbit parameters are compiled in Table 1 below

Inclination	63.4 degrees
Perigee altitude	5'784 km
Apogee altitude	41'449 km
Orbit period	14.4 hours
Repetition time	72 hours
Altitude at operation section	30'150 through 41,449 km
Operational service time	8 hours

Table 1: Orbit Parameters of LOOOPUS MOBILE D

Continuous redundant operation requires several satellites (9 in this case) relieving each other in a 8-hour cycle. Each satellite is operational in the loop near apogee, when the satellite inherent speed is lowest. This explains the relatively long operational period of 8-hr in the loop. Due to the orbit configuration, two other operational ranges are available over North America and Asia, respectively, in addition to the European zone. Figure 1 depicts the positions of the operating loops and the satellite orbits.

All nine satellites follow this same repeating ground track, that is, an observer on earth sees all the satellites in the same orbit one after the other at periodic intervals. Figure 2 shows the earth coverage zones with elevation angles of 45 degrees or more.

3. Network Architecture

LOOPUS MOBILE D is designed to be embedded into the new terrestrial system for land mobile communications in Europe which is scheduled for the nineties. This system called GSM (Group Speciale pour les Mobiles) will eventually serve 10 million subscribers throughout Europe. LOOPUS shall serve as a complementary item to the GSM network and shall expand the service area to the sea and air traffic areas

The system architecture of the terrestrial network is essentially retained since GSM and LOOPUS differ basically in the frequency used. The elements of this satellite system are presented in Figure 3, showing applications and users

As long as all 9 satellites pass along the operational part of the orbit, all satellites (of all service areas) can be operated through one central ground control station (SCC) if this station is located at an appropriate northern latitude, e.g. Norway

The same advantage applies for the Network Coordination Station (NCS). For redundancy and political reasons, however an individual NCS is foreseen for each of the sub-areas whereby two-out-of-three stations will operate in stand-by mode. Two interoperability concepts are ilustrated in Figure 4

4. Communication Capabilities

The satellite system is currently designed to provide service to about 4.000 subscribers simultaneously. Three antennas (transmit/receive/feeder link) each approximately 1 meter in diameter will be used, and a RF output of around 2 kW with a payload mass of about 250 kg will be required

The interface to the terrestrial network will be established via Mobile Switching Centers (MSC). Up to 40 MSC's are planned to be implemented per satellite system. The MSC's constitute gateways to existing national public networks, which could be either ISDN or PSTN organized Fig.5

The potential communication system architecture of one of the five useful apogee positions is shown in Fig. 6

The basic communication performance properties are as follows

- Telephone and data services (tentatively 4000 channels).
- Digital audio broadcast.
- Paging.
- Navigation (position-fixing accuracy better than 100 meters using atomic clocks on board)

5. Frequencies

Figure 7 shows the applied frequency plan. The 11/14 GHz is used for the feeder link and the 12/14 GHz band for the mobile link with an occupied bandwidth of 100 MHz each. Interference with services from the geostationary orbit is precluded because LOOPUS is in radio silence mode when its instantaneous subsatellite point is below 45 degrees latitude.

6. Antennas

The generic conflict of illuminating large service areas with sufficient high power flux densitiy have been solved with a flying spot antenna generating a single spot of about 1000

km diameter footprint. The spaceborne antenna would have an area of about one squaremeter.

An active phased array antenna combined with an intelligent on-board controller performs a flexible beam hopping in accordance with the permanently updated (demand oriented) scan pattern.

7. Repeater

The channel characteristics, of the mobile channel, taking into account the high nominal elevation angle for mobile terminals, are expected to be dominated by direct propagation components. Degradations by multipath will be less significant that those in conventional 1.5 GHz from geostationary orbits. The flying spot concept makes the traditional FDMA approach for access of many users impossible. Therefore, the proposed access procedure is TDMA correlated and synchronized with the hopping of the beam. The schematic of the payload is depicted in Fig. 8.

8. Mobile User Terminals

This type of satellite reception through mobile units at 12/14 GHz is not deemed a technological problem today. Investigations have verified (e.g. University of Paderborn FRG) that the use of electrically controlled planar antennas perform well to track the satellite. Car radios, car phones and navigation can be operated in parallel using an antenna of this type. The dimensions of the antenna development model is 30 cm x 30 cm x 1 cm. The functional block diagram of the mobile user terminal is shown in Figure 9.

The equipment related to the satellite link (i.e. the TDMA buffer coder, the TDM buffer coder and the active antenna) is interfaced to the GSM car equipment in front of the UHF RF equipment tailored to the GSM system characteristics. Selection of the terrestrial or the space link is performed by a switch on the user terminal. This switch will be activated upon the availability of terrestrial/space signals.

9. Station Keeping

Velocity corrections are necessary to maintain the proper orbit parameters. At the loop intersection point where the ascending and the decending satellite meet, the distance of the satellites shall not exceed 85 km. In comparison with the requirement of geostationary satellites, the inclined orbit requires much more orbit correction energy. For a 10-years mission, conventional orbit correction means using normal propellant, cannot be justified.

Therefore orbit corrections will normally be performed by using electrical ion thrusters. They will be actuated at times when the payload is not in operation. This design requires only a common power supply system (MBB patent pending). In this way the launch mass of the satellites can be reduced to approximately 1000 kg.

Compared to geostationary satellites, there is no transfer to the circular orbit at 36000 km height with its inherent high propellant requirements. This in turn means that the launch mass of LOOPUS is only half of that of an equivalent geostationary satellite because LOOPUS will be injected directly into its elliptical orbit.

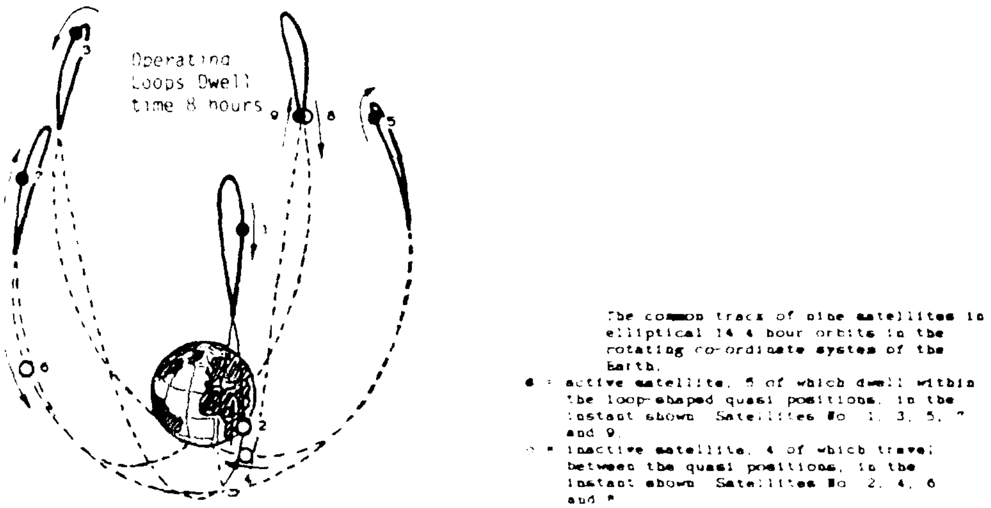


Fig 1 Common track of the 9 satellites in the rotating coordinate system of the earth

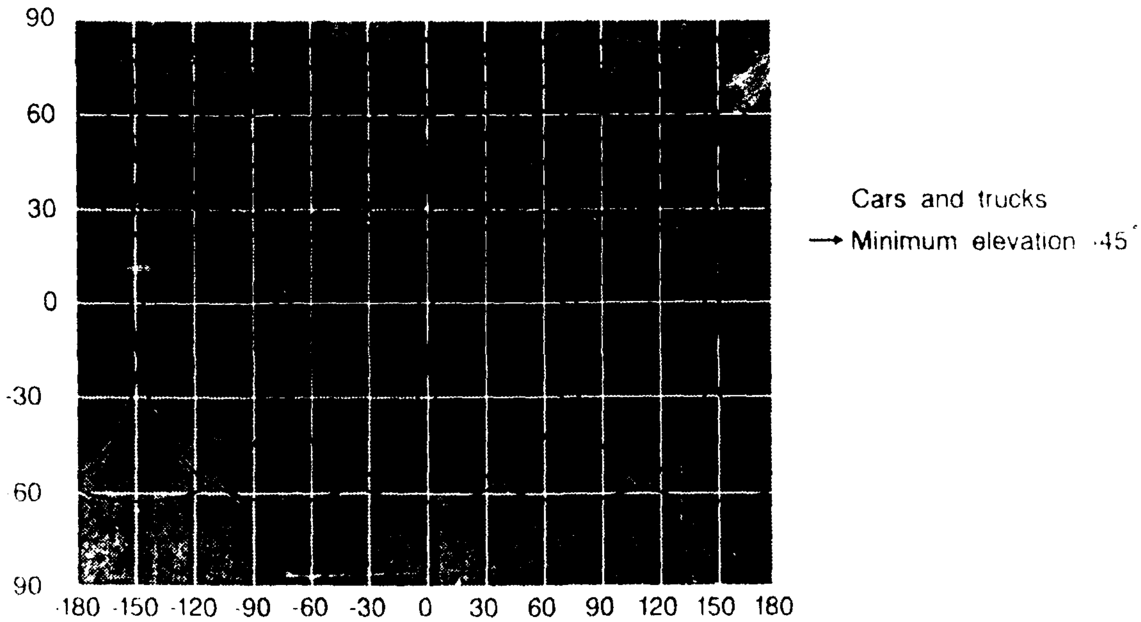


Fig.2 Coverage zones of earth station elevations of 45 degrees or more

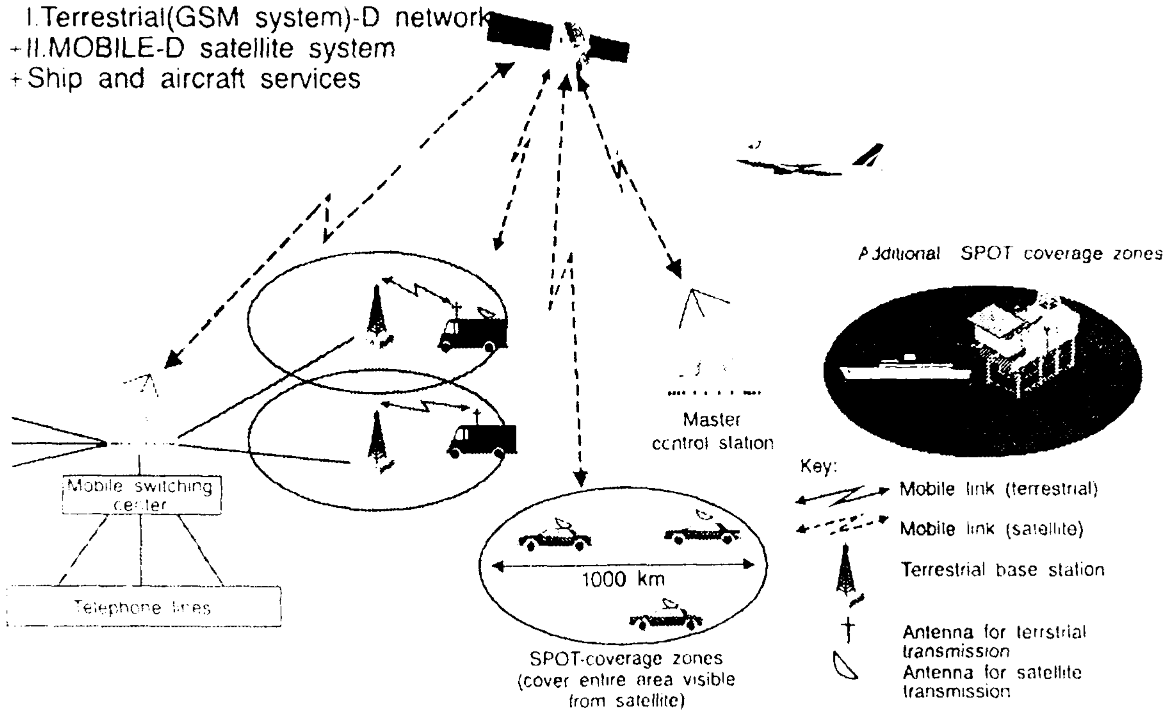
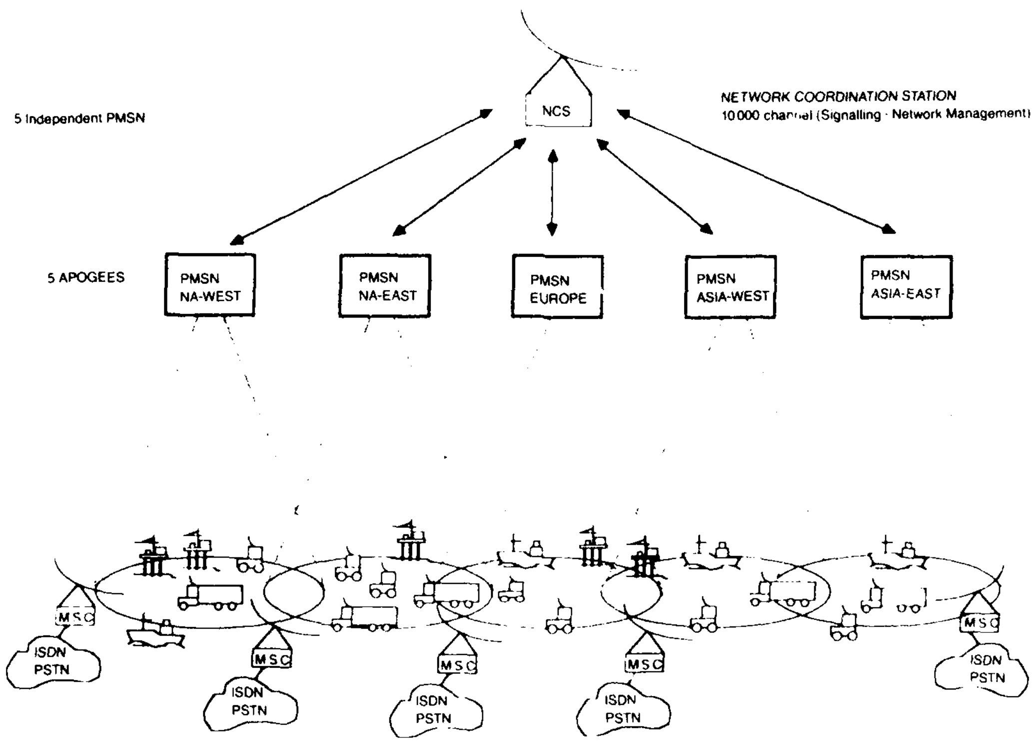
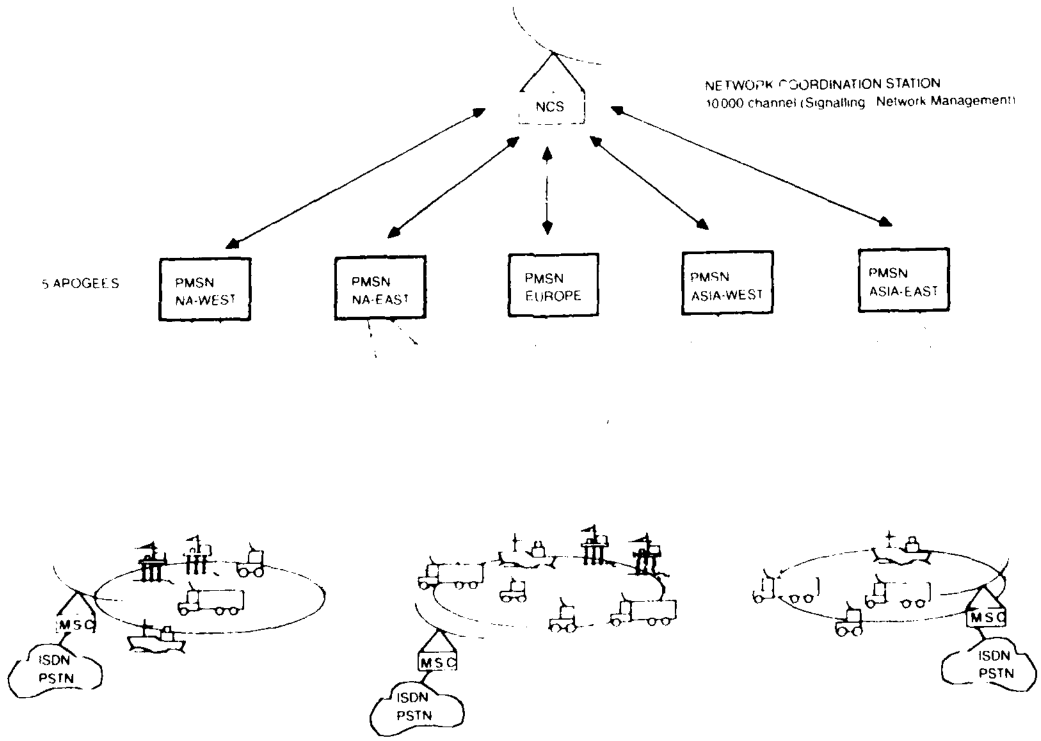


Fig 3 Joint user network architecture GSM system/LOOPUS system



(a)



(b)

Fig.4 Interoperability options for LOOPUS MOBILE D

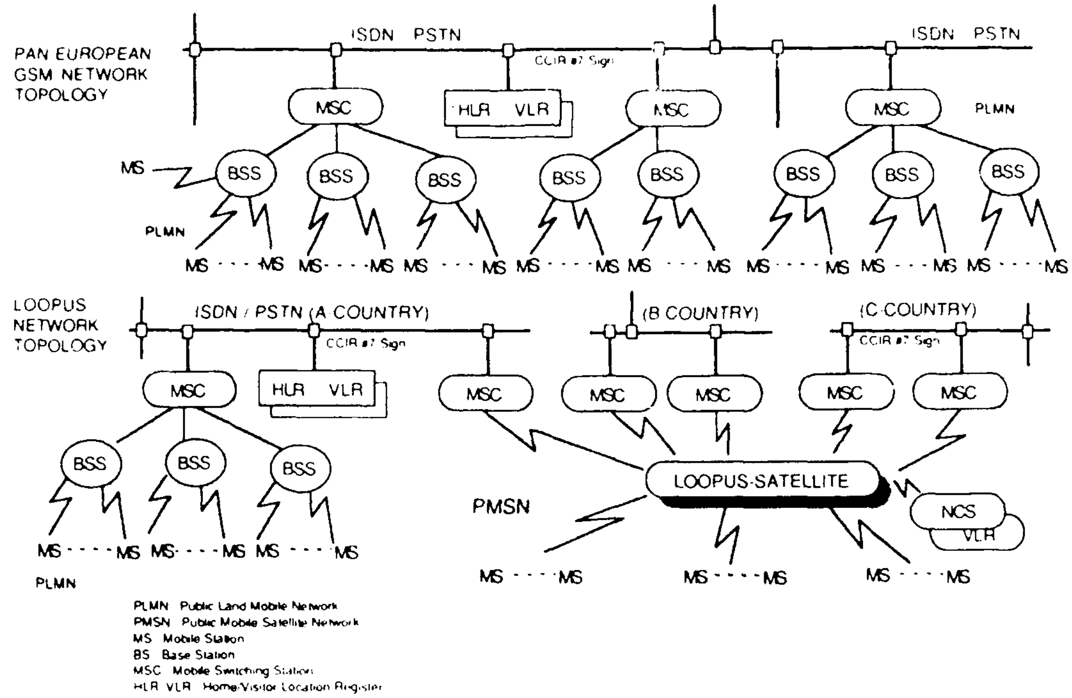


Fig.5 Comparison between LOOPUS MOBILE D and GSM network topology

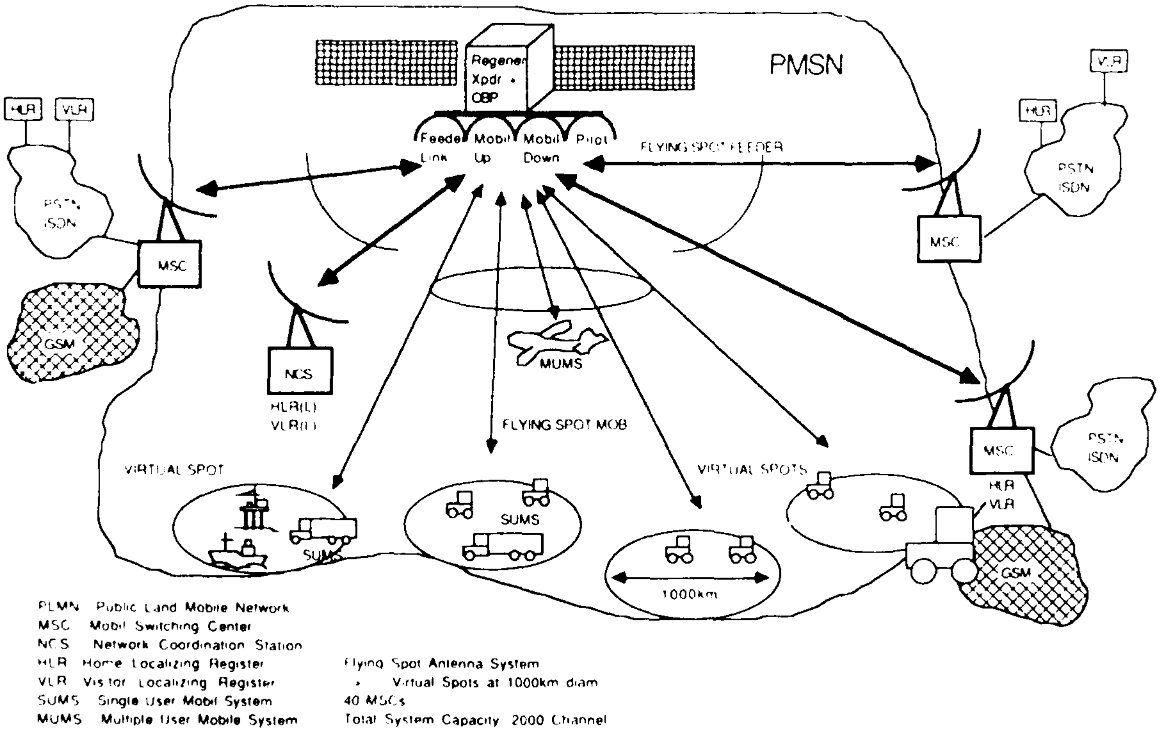


Fig.6 System architecture of one LOOPUS PMSN (public mobile satellite network)

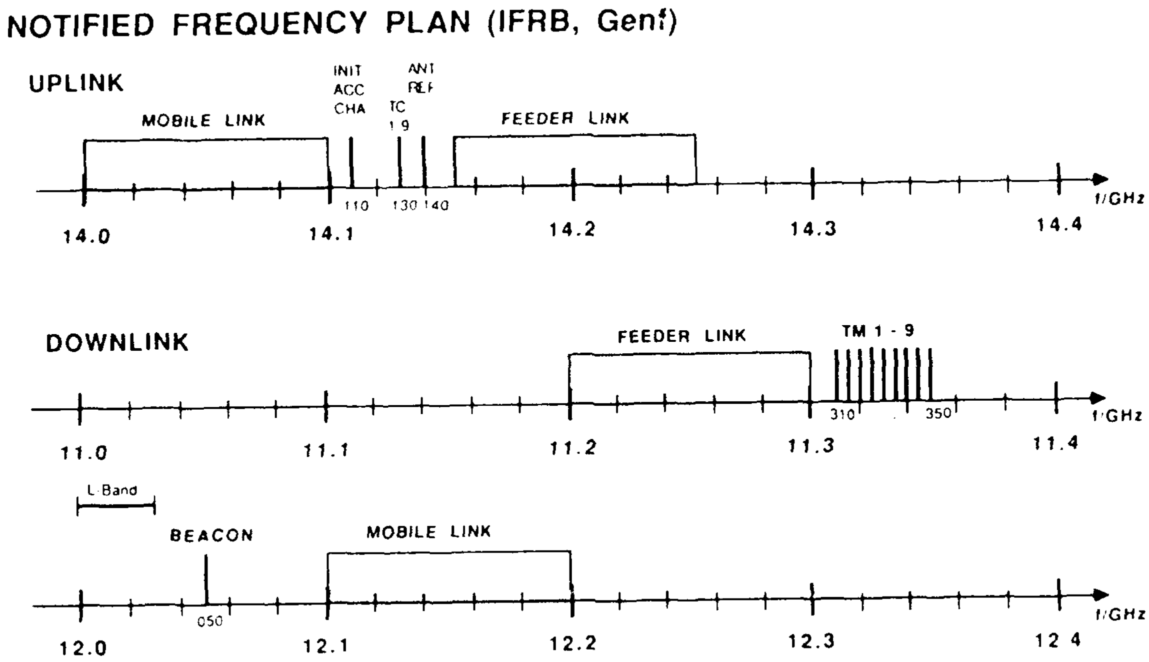


Fig.7 Frequency plan of LOOPUS MOBILE D

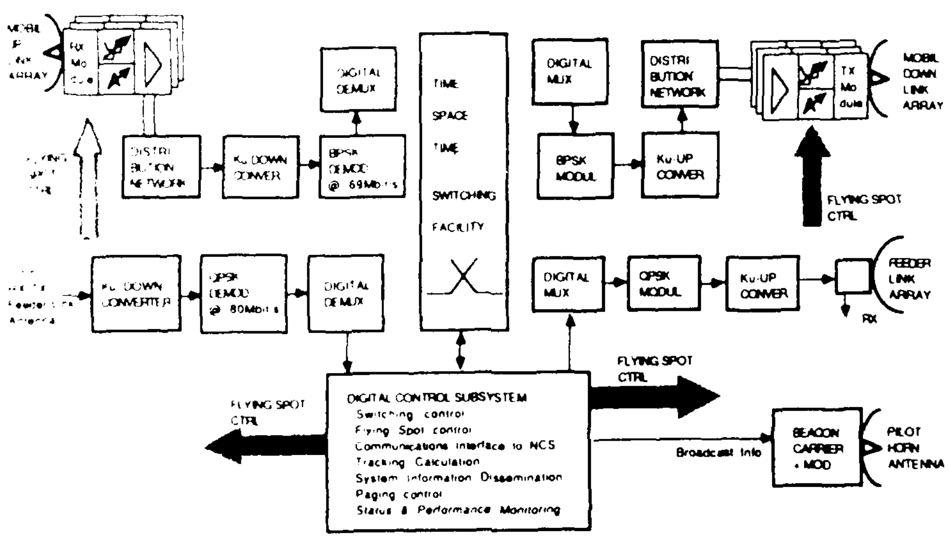


Fig.8 Schematic of the Payload for LOOPUS MOBILE D

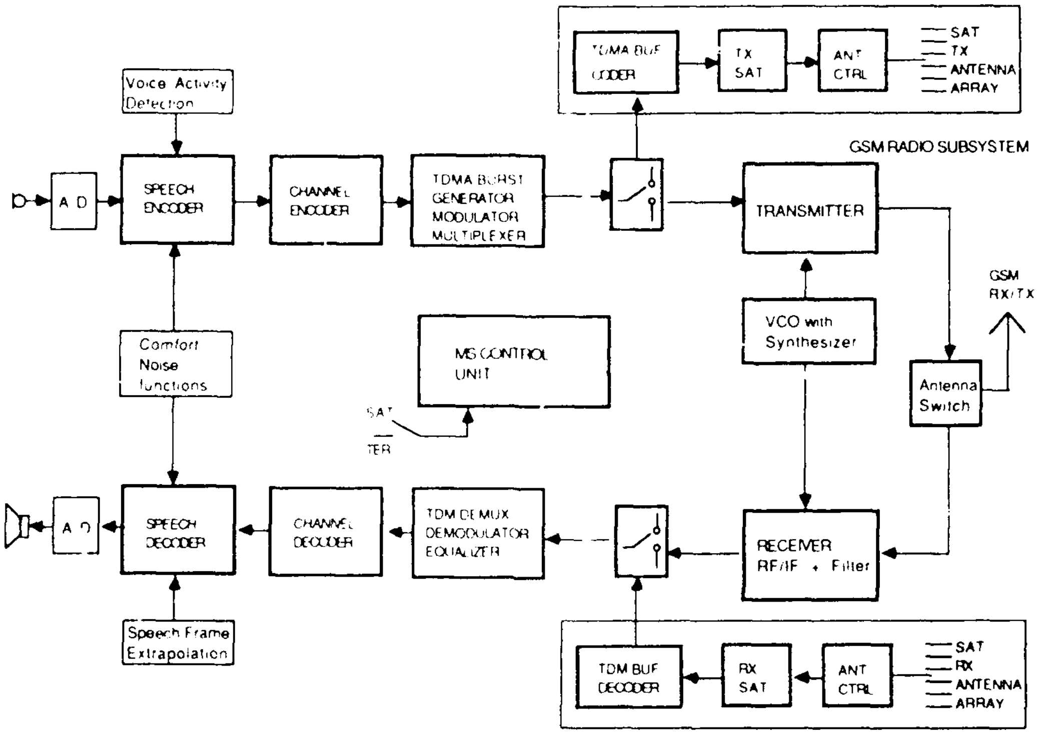


Fig.9 Block diagram of the mobile user terminal for LOOPUS and GSM

APPENDIX 9C

A REVIEW OF ESA TELECOMMUNICATION PROGRAM PLANNING

1. Introduction

In preparing for the future European space missions, ESA continues to link the mission and system studies and the technology developments, thus ensuring a strong interaction between mission and technological preparations throughout the preparatory stages of missions. The general procedure is illustrated in Figures 1 and 2.

The following outline of ESA planning will be confined to telecommunication projects. The information is extracted from ESA papers issued within the last 12 months:

- *ESA Medium-Term Plan 1988-1992*
- Long-Term European Space Plan 1987-2000
- "Preparing for the New Programmes", The ESA Technological Research and Development Programme 1988-1990

ESA and their technological research centre ESTEC are conducting advanced technology research activities, space flight missions and the implementation of adequate ground infrastructure. ESA is going to reduce their engagement in operational programmes (ECS) and will not run commercial projects.

The objectives are to support the space communication world, more and more faced with the competition of the cable world in particular optical fibres. Moreover, the competitiveness of the European industry is to be maintained and improved. What follows, is that bodies of commercial satellite communication like EUTELSAT, INMARSAT, national PTTs and even INTELSAT, take their advantage from this policy, in that they can relieve their R&D budget and obtain innovative impulses for the configuration of their future systems.

The ESA overall planning is subdivided into nine (9) themes where:

- Theme 2:Technology Infrastructure for Space Communications
- Theme 9:Common and Generic technology

are the sections most relevant to the work of the AGARD WG-13. But also Theme 1: "Earth-Space Telematics" is likely to produce some windfall products in favour of satellite communications. For more effective work and planning, but also due to budgetary cuts, ESA tries to effect coordination with other organisations.

The executive has regular contacts with other organisations involved in space technology R&D, e.g. INTELSAT, INMARSAT and EUTELSAT. The purpose of these contracts is to compare plans and to exchange views on the experience gained, to learn more about prospective technological requirments, and to adjust certain actions to make the respective programmes more complementary.

Furthermore, regular contacts with NASA serve to elaborate design guidelines and standards in new fields, and to explore areas where closer working relationship can yield benefits to both Agencies particularly with respect to in-orbit technology demonstrations in which common interest may lead to some joint flight experiments.

2. Technology Infrastructure for Space Communications

The objective of this theme is to prepare the technology infrastructure for future space communication payloads and services. The theme is justified by the evolution of the requirements of increased traffic and new services (data transmission, electronic mail, video-conferencing, etc.) and by the evolution of the technology itself, allowing more powerful and operationally flexible satellite systems through the use of digital technology, more sophisticated antenna systems, multi-beam operation, on-board re-generation and switching, frequency re-use, etc.

So far, "pilot projects" have driven the technological activities in Theme 2. This approach was extremely useful for the preparation of the ESA future telecommunication programme (PSDE) since it enabled a realistic assessment to be made of the state-of-the-art technology versus potential mission and system requrements. In the planned PSDE slice 2 programme some payloads with technology developed in Theme 2 are presently the subject of detailed phase A studies. In 1988, it will be decided which payloads will be embarked on board one of the PSDE satellites (Sat-1, -2, -3) in the 1991-1993 time frame and which will be further tested on ground for a later launch in one of the AOTS satellites due to be launched 1997 onwards.

In this context, the ASTP which addresses techniques and technology for which feasability has already been demonstrated will focus more on support to the PSDE either for the Sat-1, -2, -3 or for the AOTS payloads. The purpose of the basic TRP will be to study advanced concepts and technology for which feasability has still to be demonstrated, such as active antenna payloads

In recent years, contacts have been intensified with the EEC technology R&D programmes, i.e. RACE, BRITE, and ESPRIT and with the EUREKA programme. Due to the increasing significance of this research for space applications, the Executive intends to engage in more systematic exchanges with a view to reach a closer harmonisation of the space related R&D in the future both at the level of individual EEC development projects and at the programme policy level.

The theme subdivision defined in the past, holds true for the next planning period:

- Wideband Mission
- Narrowband Mission
- Intersatellite Links
- Reconfigurable Communication Satellites

The planning is seen as an evolutionary process and is based on the achievements until 1987.

The major technology programmes (funding sources) are listed in Table 1.

Programme name	Definition
BASIC	External contracts for the Basic TRP (Technical Research Programme)
ASTP-3	Advanced Systems and Technology Programme for telecommunications (3 stands for the 3rd planning period)
DRPP	Data Relay Preparatory Programme

Table 1: ESA/ESTEC Telecommunications Technology Programmes (Funding Sources)

The Wideband Mission sub-theme encompasses technology for future Fixed Satellite Services in Europe such as Ka-and Ku-Bands for the RF links and on-board Processing to implement a "Switchboard in the Sky" concept. These technological innovations are candidates for the ESA AOTS and are also considered by EUTELSAT for their third generation communication satellites.

The main achievements and follow-on activities are related to the following matters:

1. Baseband processor to provide a Time-Space-Time (TST) switching function. On-going work is planned in:

   - Double-hop-experiment funded through PSDE, EUTELSAT and interested national PTTs

   - ASTP-3 for technology

   - First flight opportunity with a reduced version (time switch only) on PSDE SAT-2 satellite

   - Further flight experimentation on AOTS

2. Minituarised 12 GHz QPSK modulator using MIC technology and dielectric resonator carrier generator (DRO).

3. 30 GHz low-noise front end using HEMT devices.

4. 15 GHz high stability FET local oscillator using MIC technology and DRO.

5. In preparation for future European FSS and DBS requirements, the initial design trade-off study on reconfiguratble, multiple contoured beam reflector antennas for both scenarios is nearing completion. Hardware activities are anticipated in 1988 onwards.

6. Ku-, Ka-Band dichroic subreflector

The concepts and technology studied under the narrow-band mission subtheme led to the ARAMIS payload which is presently of great interest to INMARSAT. The concept: direct radiating phased array at L-band, is also contemplated for LAMEX, a Land Mobile Satellite Service for Europe. However, carry-on of LAMEX is not very likely.

On-going activities are based on successful RF development.:

1. 20 W linear solid state amplifier using a European bipolar transistor. A new concept making much better use of the power capabilities and efficiencies of transistors, is under consideration in favour of multicarrier applications, and active antenna arrays with dynamic amplitude and phase steering.

2. Light-weight compact L-band receivers

3. High-rejection narrow band filter modules using SAW technology on quartz substrate.

4. Codec for packet data and access request message transmission, based on the use of a punctured convolutional coding scheme with soft decision Viterbi decoding and automatic repeat request.

5. Multicarrier demodulator

6. Beam-forming network and channel-to-beam switching associated with phased array multi-beam antennas.

The third pilot mission, intersatellite links (ISL) with optical technology, is presently of great interest to EUTELSAT and INTELSAT. Comprehensive development of semiconductor laser-diode technology has been initiated, focussing on critical components such as high-radiance, single-mode laser diodes, optical wavelength multiplexers, spectral transmit/receive isolators and pointing and tracking devices.

Future activities in this area will deal with optimised system designs for high-rate GEO-to-GEO optical intersatellite links, followed by breadboard development of critical technology, driven by this application, e.g. coherent optical communication, holographic beam collimation and phased-array beam steering. Increasingly attention will also be placed in the use of laser-diode pumped Nd-YAG lasers for ISL applications.

The fourth mission, reconfigurable communications satellites aims at the development of technology which will allow satellite payloads to be reconfigured under ground control from their original combination of services to others better suited to the evolving demand (during the spacecraft lifetime) for different telecommunications services. Developments included within this pilot mission are:

1. Systems concept studies and initial hardware activities for implementation of active array antennas.

2. Several components required by advanced payloads have been successfully developed, e.g. low and high power waveguide switches (R and S types) lightweight transmission (TEM) lines for beam forming networks variable power divider and phase shifters.

3. Active elements are presently in development, e.g. Solid State Power Amplifier, (SSPA) modules at 12 GHz. The demonstration of Monolithic Microwave Integrated Circuits (MMIC) already initiated with a Foundry User exercise and a Foundry Process evaluation continues with the realisation of MMIC receiver front end modules at 14/12 GHz.

4. Future reconfigurable payloads will require a large amount of local oscillator frequencies which can be set on ground command. Development activity will lead to the design and manufacture of an Agile Multiple Frequency Synthesizer in tle L and C frequency bands, suitable in particular for future FSS and MSS and also DRS (Data Relay Satellite).

5. Dual mode TEM line beam forming networks and compact feed arrays for a multiple contoured beam reflector antenna is believed to be useful for C-Band communications.

6. Advanced processes for fabrication of dual gridded reflector antennas: laser etching is likely to produce higher resolution grid structures and simpler manufacturing procedures.

7. Dual shaped reflectors can provide a complex contoured beam using a single feed horn.

The ESA/ESTEC medium-term plan foresees a series of technology development programmes which are not reflected by the above subject. Details can be drawn from an excerpt attached as Annex A to this paper.

3. Missions and Experiments

The programme consists of the following activities:

- The Payload and Spacecraft Development and Experimentaton (PSDE) programme which includes the development and the necessary in-orbit experiments to prepare for the communication related mission foreseeable in th 90s.

- The Advanced Orbital Test Satellite (AOTS) which is to demonstrate another technological step on top of the achievements of the PSDE technological missions.

- The Advanced Systems and Technology Programme (ASTP) which contributes to the technology support needed by the PSDE programme.

- A first operational Data Relay Satellite (DRS) System, as part of the future in-orbit infrastructure.

Within the PSDE two in-orbit experimental missions are foreseen:

- The technology mission which consists of experimental satellites to be launched in the 1991-1994 time period,

- The advanced Orbital Test Satellite System to be launched in the time period 1997 - 1998

The PSDE programme is divided into slices. A resolution concerning the first two slices (basic support studies and pre-development of payloads and satellite configurations) has already been approved.

The Data Relay Satellite System (DRSS) space and ground segments will provide cost-effective data relay, information transmission, telemetry, telecommand and ranging services, to the foreseen European space programmes including in particular Columbus, Ariane 5, Hermes and Advanced Earth Observation Satellites.

The DRS system will be made available in a timeframe compatible with the need of the in-orbit Infrastructure (before the mid 90s). The space segment will consist of two satellites operating in S- and Ka-band. An optical pre-operational payload is also under discussion.

A Data Relay Preparatory Programme (DRPP) has been initiated to establish and define DRS space and ground segments, investigate feasability/benefits of interoperability with other DRS systems such as the American TDRSS and initiate required technology developments.

Figure 3 shows the time plan of the relevant telecommunications programmes. The associated technology projects are not shown here. A summary description and objectives is given in Annex B of this review. (It is to be noted that the planning reflects the status of 1987. Several programmes and projects are subject to revision and have since then experienced some changes.)

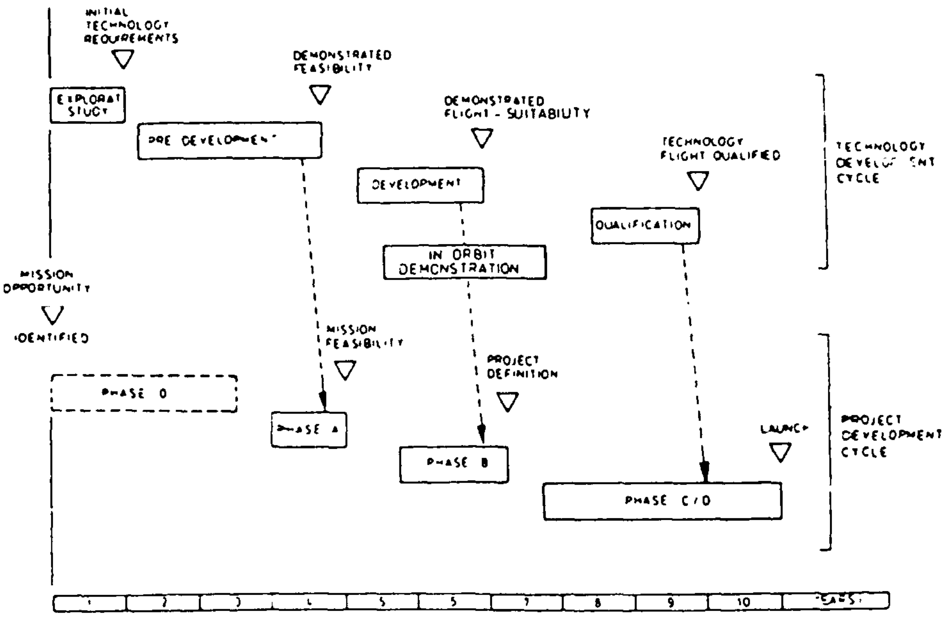


Fig.1 Technology development, in-orbit demonstration and project phasing relationship (idealised)

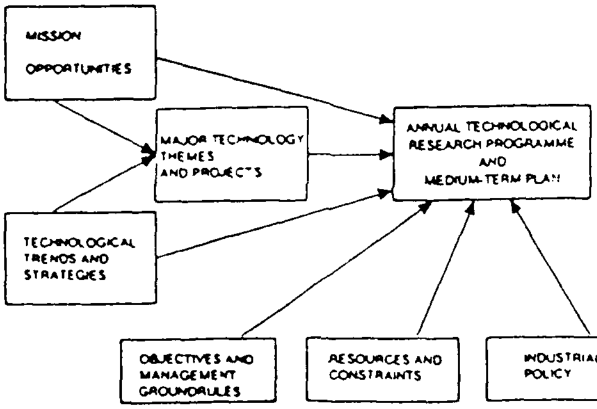


Fig.2 Planning logic

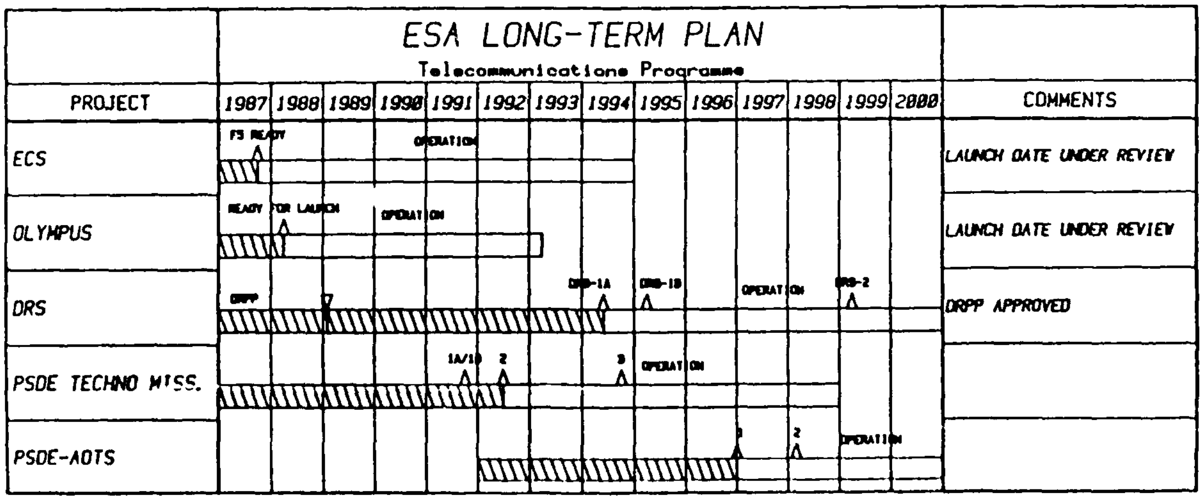


Fig.3 Time plan of the telecommunication programmes

ANNEX A TO APPENDIX 9C
ESA/ESTEC TECHNOLOGY DEVELOPMENT PROJECTS
MID - TERM PLANNING
SPACE COMMUNICATIONS INFRASTRUCTURE

1. FIXED SATELLITE SERVICES

1.1. SYSTEM

1.1.1 Coded PSK Modem

Conventional demodulators make decisions on each transmitted symbol, independent of all others. With reasonable care in signal design, the theoretical performance limit given by this method can be approached rather closely. The conventional approach to improving performance is to encode the data prior to modulation and to decode the demodulated data stream. This has the immediate disadvantage of increasing the symbol rate and hence occupied signal bandwidth and requires the demodulator to operate at an even lower carrier to noise ratio than would be indicated by the net gain in performance. The demodulation process loses information which can then not be exploited in the decoder.

The purpose of this ASTP-3 item is to develop a prototype TCM modem which makes use of a reduced-complexity decoding algorithm. Use will be made of the results of item 1.1.10.

1.1.2 Reperater Antenna Configuration For On-Board Processing

Anticipated requirements on coverage flexibility might be better fulfilled with active antennas. This will lead to a drastic change in repeater configurations. RF power generation and/or Low Noise reception will be distributed amongst a large number of elements. Critical aspects to be addressed are RF technology evaluation for amplifiers, phase shifters (with associated DC and thermal constraints), redundancy strategies in view of EIRP and G/T and frequency re-use requirements etc. Existing software tools (e.g. TOPSIM) will be used to start study but new software will have to be developed.

1.1.3 Fading Compensation Study

Studies of fading compensation techniques using up-path power control, adaptive coding and/or modulation, site diversity and frequency-band diversity, either separately or in combinations.

1.1.4 Communications Network Simulation Software

It is becoming increasingly difficult to see and treat a satillite link in isolation either from the rest of the satellite system or from the terrestrial networks to which it is connected. The complexity of the problems involved is such that computer simulation is a vital tool in the design process. The purpose of this ASTP item is the development of a simulation package which meets the requirements identified in an earlier study. It is anticipated that it will build on commercial modelling software and that it will be implemented in phases to allow a gradual accumulation of experience and expertise.

1.1.5 Coded Modulation Multibeam Satellite

Study of the ways in which the reduced sensitivity to interference afforded by integrating coding and modulation can be used to increase the efficiency of spectral re-use in a multibeam satellite system.

1.1.6 Satellite Transmission And Networking Experiment

This item involves the development of a 'pseudo-modem' to interface simulation software to a real channel, and use it for transmission studies in an 'internetworking' experiment to link various types of LAN by ECS.

1.1.7 Interactive Service Narrowband

Study, development, equipment procurement, and demonstration of interactive satellite-delivered point-to-multipoint services such as tele-education.

1.1.8 Millimeter Wave Beacon

Development of critical technology for millimeter-wave beacon payload for European coverage from a geostationary spacecraft. The payload shall transmit a triplet of coherently derived CW beacons at 48, 96 and 144 GHz, of sufficient stability and EIRP to enable reception with small ground stations. The payload will allow investigation of atmospheric characteristics in the radio windows around 90 and 140 GHz and their comparison with the 40-50 GHz band, for which beacon payloads are planned (Italsat).

These investigations are of primary importance for the exsloration of millimeter wave applications in communication and remote sensing, and the proposed study is concerned with a first detailed definition and optimisation of such a beacon payload, including the breadboarding of a laboratory set-up of the critical active items of the beacon.

1.1.9 Demodulator Small Earth Station

Study, design and implementation of a serial demodulator for offset-binary signals, for use in small earth-stations. The operating bit rate is 2 Mbps.

1.1.10 Demodulator Synchronisation Strategies For Digital Implementation

Implementation of modems with digital technology is rapidly replacing traditional analogue techniques. Particularly with VLSI developments, digital implementation offers important advantages in the areas of mass, power consumption production costs and stability. The integration of modulation with coding and other sophisticated signal processing is also pushing in this direction.

The purpose of this ASTP-3 item is to follow the previous ASTP-2 study with narrower and more detailed investigations of promissing schemes. Outline hardware designs will be produced based on real components and computer simulation will be used to assess performance.

1.1.11 Convolutional CODEC

Study, design and implementation of a low cost 2 Mbps 3/4 rate convolutional codec for small earth stations.

1.1.12 Block CODEC

Study, design and implementation of Reed-Solomon and BCH codecs suitable for application where high-rate codes are needed and for concatenated-coding applications.

1.1.13 Concatenated Coding Equipment

The concatenation of two (or more) codes is an important technique which is currently used to obtain larger coding gains that would be feasible with single codes. Present approaches are however complex and limited in their applications.

A current TPP study contract (University of Manchester) is examining the mechanisms involved in code concatenation with the aim of reducing complexity and obtaining a better match to various applications. It is likely that well-chosen concatenated codes will out perform or be less complex than equivalent single codes. The purpose of this ASTP-3 item is to build prototype concatenated coding equipment based on simple inner and outer codes and to use this in a demonstration as part of the Olympus realisation Programme (probably as part of another experiment).

1.1.14 Communication Link Simulation Software

Link (transmission) simulation is an important tool for the initial design of satellite communication systems, for the study of the effects of possible changes to the design and for the investigation of anomalies. The recently developed TOPSIM offers many features but is not a complete solution and has several major drawbacks. The purpose of this ASTP-3 item is to implement an improved simulation package.

1.1.15 System And Orbit Planning Software

1.1.16 Advanced Coding And Modulation

It has been clear for some years that channel coding yields valuable results in communication satellite systems and that modern digital technology is rapidly opening up new possibilities for low-cost, low-power implementation of coding. By careful signal design, one of the major drawbacks, for commercial communications systems, of coding, that of bandwidth expansion, can be overcome. Trellis-Coded Modulation (TCM) treats the modulation and coding together and by using a high(er) order modulation together with a (relatively) complex receiver, excellent power/bandwidth combinations can be achieved.

This ASTP item is to investigate the coice of suitable codes (study phase in 1988) and to demonstrate the performance of a concatenated-TCM scheme (hardware development in 1989).

1.2 ANTENNA SUBSYSTEM

1.2.1 Antenna Technology AT 20/30 GHz

The Phase 2 activities was completed in mid 1984. Design manufacturing and testing of an electrical breadboard of a small portion of the "City Beam" feed array and network was satisfactorily completed. The breadboard included 7 profiled circular horns and a divider network of a new (ESA patented) design for feed elements shared between beams; excellent performance was obtained in the 18.1 to 19.2 GHz range and electrical feasibilty has been demonstrated. Structural/Thermal design and partial breadboarding have been conducted for the 3.7 meter reflector with hinged tips. The breadboarding included tests on specimens to predict initial parameters such as thermal expansion, thermal conductivity of CFRP and MLI efficiency. The structural/thermal analysis demonstrated that the design is compliant with specifications in particular with the in-flight surface.

1.2.2 Dielectric Maximum Aperture Feed Antenna

1.2.3 Ku-Band Active Antenna Model

The activity is part of the work towards a complete demonstration model of a communication payload using active antennas operating at Ku-band. Following initial studies in which several candidate antenna optics (e.g. imaging systems, direct radiating arrays) are studied, a concept will be selected for the demonstration model.

The activity will comprise a detailed design and optimisation of the radiating system, design and breadboarding of the feed array, and demonstration and testing of the system as a passive antenna, before the active parts-developed under other contracts-are integrated into the feed array.

1.2.4 Antenna Arrays For Active Antenna

This activity addresses the technologies that could be anticipated for a third generation FSS. The overall systems requirements should come from studies to be undertaken either in General Studies or in the TPP. To date it is considered that re-configurability well be essential, which should lead to different payload configurations. Most likely active antennas will be considered and possibly phased arrays if the microwave technology for power amplifiers and receivers is up to the requirements.

1.2.5 Foldable Dual Offset Antenna

Design and breadboarding of critical mechanical technologies for a towerless antenna configuration which requires complex deployments and stable booms for antenna dish localisation.

1.2.6/7 C-Band Antenna Technology (A/B)

In addition to the Intelsat system, C-band systems are still adopted for many regional and domestic services. This activity is concerned with upgrading European technology, particularly in the area of beamforming networks and feed arrays, to flight standard.

The current phase I activity involves the design and critical breadboarding activities for a multiple contoured beam C-band antenna subsystem. Phase II will serve to advance this technology from breadboard to flight standard and shall address the design manufacture and test, including critical environmental testing of a complete C-band antenna (i.e. feed horns, TEM-line, beamforming network and reflector). Development and measurement of a demonstration model of complex contoured multibeam antenna.

1.2.8 Reflector Antenna Analysis & Synthesis

The GRASP and COBRA software packages are ESA's standard computer programs for the analysis of reflector antenna systems and the synthesis of multifeed contour beam antennas. This new activity will address the research into and implementation of new concepts in analysis and synthesis of complex reflector antenna systems. Faster analysis techniques of the use of totally optical methods for reflector analysis. In the area of synthesis, faster optimisation schemes shall be addressed. The synthesis software COBRA shall be extended to dual mode, reconfigurable contoured beam antennas and shall include the synthesis of the beamforming network in the loop.

1.2.9/10/11 Polarisation & Dichroic Surfaces(A/B/C)

Development of a laser cutting technique and manufacturing of polarisation sensitive surface PSS and dichroic surfaces by this technique.

Laser cutting process will give a large flexibility concerning

dichroic patch design, polarisation wire grid dimensioning and is a useful technique of edging dichroic sunshields. Intermediate layers of kaptofoil embedded in sandwich structure will no longer be necessary. Resulting problems concerning mechanical integrity are no longer applicable.

1.2.12 11/14 GHz Large Off-Set Antenna

The hardware development for this reflector was performed under ASTP-2 item P 1.8 (ESA/IPC (82) 27, 56 th IPC). The programme included environmental testing of the hardware. The purpose of this extension is to conduct and evaluate the results of environmental tests. Acoustic and solar simulation with distortion measurement tests and test predictions to verify and qualify the large 11/14 GHz reflector.

1.2.13/14 Off-Set Unfurlable Antenna

The objective of this antenna is to demonstrate the technological capability to design and manufacture unfurlable antennas in the range of 4-10 m. Earlier phases selected a double folded rib concept. In the middle of 1985 the programme was modified and extended to ensure demonstration of a 5 m dish capability compatible with M-SAT requirement. As a result of completed funded activities at the end of Phase 3A all components of the TD and integration facilities will be available. Phase 3B is required to assemble, integrate and verify the TDM.

1.2.15 Reconfigurable Reflector Studies

Investigation of the potential of in-orbit reconfiguration of reflector surfaces both for dynamic correction of the profiles of large deployable reflectors and as a means to change the shape of reflectors to reconfigure coverage zones. The study shall investigate the application and mechanical realisation of these concepts with mesh, petal and possibly inflatable antennas. This should form a preparatory phase to hardware realisation

1.2.16 Development Of Dichroic Subreflector For Multifraquency ANTENNATesting

Upgrading of planned manufacturing test. Sample subreflector to electrical breadboard level, electrical testing, and manufacture of a second model for environment testing, as well as environmental tests.

1.3. RF FRONT-END

1.3.1 Low Noise Fet Front End

The development of an engineering model of a low noise 30 GHz FET amplifier having a target noise figure of 3.5 to 4 dB over the 27.5 to 30 GHz frequency band. This activity is a follow on from a previous development of a breadboard amplifier having a noise figure of 5.5 to 6 dB over the same band.

1.3.2. Gaas Synchr. High Speed Data Transmission

The utilisation of the advantages of GaAs technology to fulfil the new performance requirements for high speed data transmission will be investigated. For this purpose the circuitry for synchronisation and decommutation of high speed data stream forming standard building blocks on-board and in associated high speed ground segments will be implemented and tested.

1.3.3 Development Advanced Local Oscillator

Design, development and test of a synthesised local oscillator. This technology is applicable to future Ka and Ku band communications satellites.

1.3.4 12 GHz Solid State Amplifier

Development of power amplifiers of high efficiency and linearity.

1.3.5 RF Filters

Development of new filters for improved performances and array applications.

1.3.6 Hybrid 12 GHz Low Noise Amplifier

1.3.7 Development Of Power Divider 12 GHz

1.3.8 RF Switching Matrix

1.3.9 Variable Level 20 GHz TWT

1.3.10 Fast Power Switch For 20 GHz, Beam-Forming Network

In view of applications in active antenna devices, switches need to be developed for the BFN. The high speed power switch will be breadboarded at 20 GHz. Other components like Variable Power Dividers (VPD's) and Variable Phase Shifters (VPS's) will be manufactured in a later stage. Development/evaluation of waveguide fast 20-100 (microsecond) switch for transmit signal routing

1.3.11 2MB/S Demodulator With Fast ALC

For coherent QPSK demodulators the performance in terms of BER is still level dependant in contrast to differential demodulators where recent developments have shown capabilities to cope with 20 dB dynamic range.

1.3.12 Development Seamless Twist Waveguide

1.3.13 Frontend Receiver Multicarrier Signal

Design, development and test of a front end receiver to down convert and channelise low bit rate QPSK Modulated multicarrier signals.

1.3.14 Linearisation TDMA Repeater

1.3.15 Flight Demonstration 50 W TWTA

The opportunity exists to fly a number of European Travelling Wave Tube Amplifier (TWTAs) on F2 and/or F3 of Eutelsat II

Such a flight is necessary to demonstrate the competitiveness of European TWTAs and to prepare for commercial openings in future European and Worldwide FSS (Fixed Service Satellites) markets. This activity is for the production, test, delivery and evaluation of these TWTAs

1.3.16 TWTA Technology Improvements

1.3.17 12 Channel Contiguous MUX

1.3.18 Demonstration Of Local Oscillator Technology

The activity proposed will cover two distinct parts as below,1.Technology demonstration (hybrid line qualification). 2. Circuit Design/Hardware development (dividing phase lock loop, X-K band multipliers).

For both parts a design phase I will be started during the 4th quarter of 1987, followed by a development phase 2 lasting from early 1988 to mid-1989.

1.3.19 12 GHz Power FET Amplifier

1.4 Signal Processor

1.4.1 Wideband Processors/CAD Tools

Implementation of "integrated" Computer Aided Design tools allowing specification, test and production, in a cost effective

manner, of the customisation of the Wide Band Processor. The objective is to optimise tha decomposition of the Processor in chips and to reduce cost of the full customisation to the level of semi-customisation (gate arrays....) while increasing by an order of magnitude the level of integration level obtained.

1.4.2 Improvement Of TST/SS-TDMA System

1.4.3 FET Modulator Design (MMIC)

Present designs (at least in Europe) are based on pin diode implementations. A design using dual gate FETs should be evaluated, especially for two reasons: the data rate may be in the future 600 Mbps (Columbus project), pin diodes have limited switching speeds. In the future the modulator could be integrated with carrier generators, which would be a dielectric stablised FET oscillator. A microwave monolithic integrated circuit design of the two units could be conceived in the future.

1.4.4 Wideband Processor

1.4.5 On Board Processor TST/SS-TDMA

Development of processor using semi-custom Integrated Circuits. Future application to AOTS and mobile satellite services.

1.4.6 Multipurpose TDMA/DSI

Development of TDMA/DSI Traffic Terminal for TST/SS-TDMA System and EUTELSAT/INTELSAT TDMA System

1.4.7 STP For TST/SS-TDMA Master Station

Design and development of the hardware and software tools to implement: Variable Origin function, Multipoint Videoconferance Digital Exchange traffic routing, on-board clock generator with doppler compensation.

1.4.8 Alarm Subsystem For TST/SS-TDMA

The purpose of this study is to define the alarm, control and monitoring functions for the baseband part of the onboard TST/SS-TDMA system (item 1.4.1)

1.4.9 2MB/S MCD (Custom Implementation)

Development of multi-carrier demodulator (MCD) for TST/SS-TDMA experiment. Breadboard design and development using semi-custom integrated circuitry.

1.4.10 Test And Experiment Of A TST/SS-TDMA System, Modems For The "Double Hop Experiment".

The Agency has developed, at laboratory standard, the basic elements of a TST/SS-TDMA system, that is characterised by on-board processing and switching and therefore ...ntated to interconnect a large number of small earth statior... ...ered via high gain spot beam antennas.

A first step is the so-called Double Hop Experiment. The Double Hop Experiment and the hardware required will be studied and defined in detail in a contract under PSDE-Slice 1 which is currently being negotiated. The development and procurement of the modems to be installed in PTT Earth Stations is the most urgent activity in view of the limited availability of EUTELSAT transponder capacity. These modems are the subject of this procurement proposal as part of item F 6.8 included in the ASTP-3 work plan.

1.4.11 2MB/S On Board Viterbi Decoder

Follow-on from ground development to flight technology, for use in fixed satellite service using regenerative transponder.

1.4.12 Baseband Package For ISL

Breadboard development of package for Inter-Satellite Link, followed by Electrical Model development using semi-custom integrated circuits.

1.4.13 TST/SS-TDMA System Specification

1.5 MISCELLANEOUS

1.5.1 Receive Earth Station For Apollo

1.5.2 Earth Station Test Equipment

Phase 1: Identify European sources for microwave and RF test equipment compatible with automatic measurement systems.

Phase 2: Compose and publish a catalogue with the results of phase 1.

1.5.3 Up-& Down Converters 30 GHz Station

Study and development of up and down converters suitable for use in earth stations operating with 20/30 GHz payloads

1.5.4 Calibration Techniques For Large Antennas

1.5.5 Electronic Beam Steering Microwave

Study and development of a feed system capable of electronic tracking over small ranges

1.5.6 Front Ends For Advanced Earth Station

The arrival of FET technology entails a revision of the classical concepts used for the construction of earth stations. Most electronics may be integrated in the feed, possibly located in the prime focus of the main reflector. Study and development aiming at integrating the antenna/LNA/HPA to a maximum extent

1.5.7 Radiometers

The use of the higher frequency bands entails the need of radiometers to evaluate propagation conditions in real-time during IOT measurements. Suitable instruments seem not yet available. Devlopment of a suitable 20/30 GHz radiometer

1.5.8 TWTA For 30 GHz

Design and development of medium and high power travelling wave tub amplifiers for use as drivers or final amplifiers in earth station transmitters

1.5.9 Solid State Amplifier For 30 GHz

Design and development of solid state amplifiers for use at frequencies around 20 and 30 GHz. This comprises low noise amplifiers and high power amplifiers.

1.5.10 Passive Microwave Components 30 GHz

Design and development of passive microwave components for frequency bands around 20 and 30 GHz.

1.5.11 Small Terminal For Developing Countries

Development of 4 m diameter remote earth terminal using integrated European expertise and hardware

1.5.12 Study Of Digital Beacon Receiver

1.5.13 2 MB/S TDMA Traffic Terminal

Development of Traffic Terminal (TT) for TST/SS-TDMA

experiment. Design and development of TT with testing of other elements of TST/SS-TDMA.

1.6 ADVANCED COMPONENTS

1.6.1 Propagation Modelling For Dual Polar Radar

1.6.2 Radar Studies For 20/30 GHz Propagation

A study of the propagation properties of rain, melting band and ice, using the FM-CW radar of T.U. Delft. This equipment allows Doppler determination of fall speeds, and dual polarisation measurements using a novel type of polariser. The unique combination of these two features is of great interest for a study of propagation models for rain and melting band, in particular for application of frequencies above 15 GHz.

1.6.3 Advanced Modulation Techniques

The future trend in digital modulatin techniques is to combine for optimum performance coding and modulation. Compared to conventional QPSK a coding gain of about 3dB can be achieved without compromising on bandwidth.

1.6.4 Dielectric Resonator Filter Technology

The dielectric resonator approach will also be used at low power in shaping filters in conjunction with 12 GHz modulators/demodulators. The work will be aimed at proof of concept bread-boarding. If successful, follow-on development activities are envisaged to engineering model level.

1.6.5 Technologies For Low Noise Amplification

High Electron Mobility Transistors (HEMT) have demonstrated superior low noise performances from 18 to 60 GHz. The first investigation will cover half micron gate HEMTS to be tested for application in 30 GHz low noise amplifiers in replacement of quarter micron FETs. The second study will concentrate on quarter micron HEMTS, when they become commercially available, in order to reduce further 30 GHz amplifier noise figure.

1.6.6 Miniaturised Solid State Amplifier Technology

Investigation of competitive techniques for reduction of size and weight in communication payloads, using monolithic (MMIC) or miniaturized hybrid technologies (MHIC). The feasibility of MMIC amplifiers including their DC supply, thermal control and command circuits will be demonstrated (standard building block). Other types of MMIC components, e.g. mixers, oscillators, phase shifters, switches, required in advanced payloads, will be assessed at breadboard level. Technology evaluation will be performed in 1987. Proof of concept demonstrations are foreseen in 1988 and 1989.

1.6.7 Interference By Rain Scatter

Study of the structure of rain cells and simulation of interference configurations in millimeter wave communication systems, using available radar data. Development of models for scatter of millimeter waves by rain for application to the rain data. Development of a model which enables prediction of the amount of coupling in interference paths from a knowledge of rainfall footprint intensity statistics.

1.6.8 Advanced Optical Technology For Telecommunication Payload

Steady progress in optical technologies make it possible to consider them for implementation in telecommunication payloads. The aim is to develop a proof-of-concept model of a optical beam forming network including all the constituent components, integrate it within an existing RF payload and perform and overall evaluation test.

1.6.9 Millimeter-Wave Beacon Payload

Development of critical technology for millimeter-wave beacon payload for European coverage from a geostationary spacecraft. The payload shall transmit a triplet of coherently derived CW beacons at approximately, 48, 96 and 144 GHz of sufficient stability and EIRP to enable reception with small ground stations.

1.7 INTER SATELLITES

1.7.1. 1-1.3 Micron Receiver

The 1.3 micron spectral region holds promise for laser diode ISL applications because of the reliability and power output capability of laser diodes operating at this wavelength. Yet the sensitivity of current detectors in this spectral region is generally insufficient for the envisaged application. The proposed activity will therefore focus on the development of appropriate detector technology, with special emphasis being put on achieving 4-quadrant devices.

1.7.2 Diode Pumped Nd-Yag Laser Transmitter

This activity will deal with the development of a diode-pumped Nd-YAG laser for potential application in inter-satellite communications.

1.7.3 Laser Diode GEO-GEO ISL Package

In a first stage, this activity will deal with the design of an ISL optical communication package to be used as a reference guide for subsequent technology development. The system will be defined against the requirements of a high-rate, duplex GEO to GEO intersatellite link. In a second stage, the activity will deal with the breadboard development of critical technologies driven by this application, such as: coherent optical communication as a possible means to improve link performance in the presence of solar background; advanced optical techniques such as holographic collimating optics; and phased-array optical beam steering arrangements to improve system performance while reducing overall payload mass, size and power.

1.7.4 Optimised High Power Laser Diode

This activity is concerned with the development of high power semiconductor laser sources for intersatellite optical communication systems. Based on theoretical studies and exploratory research work conducted under TRP funding the development of representative devices is now proposed under ASTP-3. Activities will initially focus on achieving reliable operation and single lobe emission with diode arrays at power levels in the order of 200 mW. In continuation, work will concentrate on achieving higher output levels, high frequency stability and lifetime.

1.7.5 Optical Nd-Yag & Vibronic Laser

This activity will be concerned with the development of high-speed modulator technology for laser sources.

1.7.6 Optical Technology High Speed

1.7.7 Power Amplifier For Millimeter Waves

1.7.8 Parametric Amplifier For Millimeter Waves

1.7.9 Optimum 1.3-1.7 Micron Laser Diode

This activity is concerned with the development of high power semiconductor laser sources for intersatellite optical communication systems. Based on theoretical studies and exploratory research work conducted under TRP funding, the development of representative devices is now proposed under ASTP-3. Activities will initially focus on achieving reliable

The scope of this development is a demonstration model of the baseband processor fc; a regenerative payload. In the forward direction the digital).ocessor will distribute the up-link regenerated TDMA signal into a number of TDM down-link signals. In the return direction forward error correction, bit stuffing and assembling of the up-link SCPC signals into one single TDM down-link will be performed. Access control, ARQ and house.'keeping fun' 'ions must also be performed.

2.4.2 Forward Link Test Transmitter

Follow-or work from proof-of-concept studies that will develop a hardware design.

2.4.3 Return Link Test Receiver

2.4.4 Regenerative Repeater Technology

2.4.5 Forward Feederlink Regenerator

Follow-or work from proof-of-concept studies that will develop a hardware design.

2.4.6 Multicarrier Modulator

Follow-or work from proof-of-concept studies that will develop a hardware design.

2.4.7 Return Feederlink Multiplexer

Follow-or work from proof-of-concept studies that will develop a hardware design

2.4.8 On-Board Switching & Control Subsystems

Follow-or work from proof-of-concept studies that will develop a hardware design

2.5 MISCELLANEOUS

2.5.1 Algorithm And Data Transmission, Navsat

2.5.1.1 Radiation Hard Elliptical Navsat Orbit

The particle environment experienced by a Molniye type orbiting satellite is worse that for the GEO satellites. This is to study how to obtain a 10 year lifetime for a satellite in Molniya type orbit.

2.5.3 NAVSAT Ground Control

A follow on to an ASTP-2. activity to address the ground control aspects of NAVSAT.

2.5.4 NAVSAT Antenna System

The NAVSAT spacecraft may in addition to the nevigation payload also comprise a communications payload. The antenna system to study should therefore comprise an L-band antenna for up and downlink and a C-band antenna for up link. The antenna patterns should be matched such that each location on the ground receives about the same EIRP. A compromise of pattern and distance dependance should be reached.

2.5.5 Navigation Receiver For NAVSAT

The NAVSAT system will transmit two signals, a PRN-code and a CW signal, for navigation. Two classes of receivers are therefore foreseen and have to be studied in some detail before breadboarding.

2.5.6 Pulsed Transponder For NAVSAT

To develop a power efficient L-band transponder for the NAVSAT TDMA signals.

2.5.7 Navigation Processor

Predevelopment of a processor, with navigation S/W to measure range, handle Doppler adjustments and perform system time management. It is envisaged that such a module will form part of a terminal within the Satellite Navigation Experiment. The development is needed to achieve a full system test.

2.5.8 Electronically Steerable Terminal Antenna

Multipath reception is the major source of interference with low G/T terminals anticipated to be used for future Mobile-Satellite Services. The study shall analyse whether suitable signal and beam processing can be implemented in an economically attractive way to overcome these impairments.

2.5.9 Small Electronically Steerable Landmobile Terminal

3 DIRECT BROADCAST SERVICES

3.1 SYSTEMS

3.1.1 Reconfigurable Multi-Mode Feed Networks

Design manufacture and test of an in-orbit reconfigurable multifeed antenna including RF sensing.

3.1.2 Advanced Payload Configuration, Active Phased Array

3.2 ANTENNAS

3.2.1 Reconfigurable Contoured Antenna AT 3IW

Design, manufacture and test of in-orbit reconfigurable multi-feed antenna with application to DBS missions. RF sensing will also be performed.

3.2.2 Multicoverage Antenna AT 3IW

Shaped reflectors can be an attractive solution for contoured beam coverage requirements both for fixed satellite services and direct broadcast satellites. Having addressed the design and scaled breadboard testing of a single feed dual shaped reflector antenna for contoured beam coverage of Europe, the follow-on activities shall address design and manufacture of a feed network for the existing antenna developed from previous ASTP funding. Refurbishment of existing antenna an AIT of complete antenna as demonstration model. Upgrading of antenna test facilities for multi-coverage antenna

3.2.4/5 Shaped Dual Reflector Antenna (A/B)

Design manufacturing and testing of a demanstration model dual reflector antenna for contoured beam shaping. RF sensing of beam will be performed

3.3 RF FRONT END

3.3.1 R-Type Microwave Redundancy Switch

Six-port 3-position waveguide switch development for redundancy-switching applications in the transmit section of satellite communications payloads. Particular requirements: low mass, low insertion loss, high isolation, high reliability. Study and hardware feasibility demonstration to be followed by EM development.

ANNEX 6 TO APPENDIX 9C
ESA TELECOMMUNICATION
LONG - TERM PLANNING

TELECOMMUNICATION

1. BACKGROUND CONSIDERATIONS

The achievements in this field over the past ten years are particularly significant. ESA has carried out the OTS, MARECS, ECS projects, a significant European industrial capability in space communications has been built up; Europe is involved in Inmarsat; Eutelsat has been established; the next demonstration mission, OLYMPUS, is being prepared and European efforts in the study, development and demonstration of advanced systems and related technology are proceeding.

Future evolution in this field and the capability of Europe to play a mojor role will depend on a complex mixture of various factors such as:

- The market for communication satellites is rapidly expanding with the economic impact of space communications going well beyond the space segment alone, into the earth segment and terminal markets.

- The introduction of high capacity optical fibres and the digitalisation of the terrestrial network will be the major events of the telecommunication evolution in the next three decades. This will affect both regional and inter-continental communications. In Europe, following ISDN, an Integrated Broadband Communication Network (IBCN) is being prepared for implementation in the second half of the next decade on the initiative of the European Community through the RACE programme. Future satellite systems could be integrated advantageously in this network, provided that new dechniques, improving the flexibility of the space segment, such as on board processing, are developed in time.

- New satellite systems are being and will be introduced in the next decade such as those for land and aeronautical mobile communications, video conferencing, high definition and pan-European television, navigation and position determination. This will happen in a competitive and, in Europe, progressively less regulated environment.

- Other space missions and applications and in particular the various elements of the In Orbit Infrastructure will need telecommunication support in order to increase their efficiency and reliability. This enormous task should not be performed only by the U.S. TDRSS network. A European Data Relay satellite system, possibly integrated with US and Japanese equivalent ones, will clearly become necessary.

- North American companies can make considerable investments in research and development, amortised over a large and rapidly developing domestic satellite market which will account for more than half of the world market in the period 1986-89.

- Large amount public funds in the US are invested in advanced communication satellite techniques. For instance, twenty-five billion US$ will be invested in the next few years for the feasibility studies of the Strategic Defense Initiative (SDI) programme. It is estimated that at least a third of this investment is directly or indirectly related to advanced space telecommunications and informatics. The Department of Defence (DOD), with the 10 billion US$ MILSTAR project, and NASA, with the Advanced Communication Technology Satellite (CTS), will develop fundamental techniques in US industry that will be applied to the commercial communication satellites of the next decade.

- European companies have increasing difficulties in obtaining transfers of technology from partner companies in US.

- Japan, which is extremely active in the ground station area, also has a large programme of experimental and operational satellites for the 90's and later, in practically every area of space communications including domestic mobile, 20/30 GHz high capacity, millimeter wave, data relay and optical inter-satellite communications.

- Commercial operators such as INTELSAT, INMARSAT, EUTELSAT, etc. cannot make large investments in advanced technology. Systematic demonstration of new hardware techniques is not their role.

- Separate national technological developments and promotional flights often lead to a duplication of industrial efforts which is clearly incompatible with European limited resources and with the costnatly increasing international competition.

2. OBJECTIVES OF THE PROGRAMME

The above considerations show that a strong, coherent and unified Agency effort in space communication is needed to ensure that European industry can maintain and expand its competitive position in the space communications market, and that Europe can operate with the necessary efficiency and independence in other space fields. This effort should have the following objectives:

a) To develop and ultimately test in orbit specific advanced space techniques which will contribute to the long term development of established communications systems.

b) To demonstrate and promote new space communication services for the expansion of European space communication activities and the development of a larger domestic commercial market, as the necessary prerequisite of any successful export effort.

c) To support other space missions and applications, through the development of a European in-orbit communications infrastructure (Data Relay), as an integral part of the general in-orbit infrastructure (IOI).

3. PROGRAMME SCENARIO AND CONTENT

The telecommunication programme consists of:

- The Payload and Spacecraft Development and Experimentation (PSDE) programme which includes the developments and the necessary in-orbit experiments to prepare for the communication related missions foreseeable in the 90's and beyond.

- The Advanced Systems and Technology Programme (ASTP), which contributes to the technology support

needed by the PSDE programme.

- An operational data relay satellite (DRS) system, as part of the future in-orbit infrastructure.

The development of a demonstration/preoperational navigation system is an additional option of the programme.

The schedule of future ESA telecommunication satellites is given in Figure 1.

3.1 The PSDE programme

The PSDE programme is the backbone of the future ESA activities in the field of telecommunications as, in a unique framework, it includes the studies such as those of the type formerly undertaken by the Telecommunications Preparatory Programme (TPP), the developments and the necessary in-orbit experiments to prepare for the communications related missions foreseeable in the 90's.

The general objectives of the programme are to:

- Improve the competitiveness of European industry in the communication market of the 90's
- Experiment with and demonstrate new services to enlarge the field of applications of satellites.
- Introduce and experiment with new techniques for designing satellite systems more competitive from the economic point of view with the expanding innovative terrestrial network.
- Prepare for the specific needs of operating agencies and influence the technical realisation of their future networks with experimental/pilot missions in order to ensure a position of strength for European industry.
- Identify, introduce and experiment in-orbit with very advanced techniques and technologies that, albeit promising, clearly require a convincing in-orbit demonstration before they can be considered in operational systems.

In order to meet these general objectives, a number of payloads will be defined and developed and a selected set of them may be flown as follows:

1) A Technology Mission, in the 1991-1994 timeframe,

2) AOTS, in the second half of the 90's.

Experiments and demonstrations with existing satellites, and other activities to improve the competitiveness of European products, will also be performed. This includes, in particular, development work towards platforms adapted to Ariane V launchers.

The activities foreseen in the PSDE programme are grouped in the following programme lines:

I Basic Support Line (BSL)

II Pre-development/Development of Payloads (DOP)

III Flight Configuration Studies and In-orbit Experimentation (CIE)

IV Experiments and demonstration with existing satellites (EDES)

V Improvement of Competitiveness

The PSDE programme will be executed in slices, as shown in Figure 2. These slices represent packages of activities to be sequentially defined in detail, approved and executed. Each slice is related to one or more programme lines. In principle separate declarations would be adopted for each slice of the programme.

I Basic Support Line (BSL)

This is a basic element of PSDE. The effort on this activity would be essentially constant throughout the PSDE programme.

The porticular objectives of the Basic Support Line programme are to:

1) Perform mission, system and general configuration studies and carry-out other activities with the view to preparing and continuously updating the medium and long-term programme of ESA in the field of Telecommunications.

2) Define the technology developments necessary for the execution of the telecommunication programme and their timing and provide these requirements as inputs for the technology programmes definition.

The specific tasks of this programme line are listed hereunder.

- Perform prospective studies, market surveys and economic analyses in order to identify future mission requirements and their timing in the fields of Telecommunications and space communication infrastructures such as Data Relay systems and Navigation.
- From the projected mission requirements and the advances in technology identify the need of system and hardware development and in orbit experimentations.
- Perform the mission and system initial studies that will be required in support of the definition of payload and spacecraft elements considered for development throughout the PSDE programme.
- Promote existing and future ESA experimental systems and maintain contacts with the potential users; perform studies and establish the detailed plans for using the experimental in-orbit communications capacity.
- Prepare for future experimental missions and define new programme elements when required.
- Support telecommunication operating agencies (e.g. EUTELSAT) and European PTT working group whenever necessary.
- Contribute to the international planning and regulation effort (ITU, CCIR) with special care to frequency allocation.
- Provide telecommunication support to other disciplines (e.g. Columbus programme and associated projects).
- Support the RACE programme in defining the role of satellites in the Integrated Broadband Communication Network (IBCN). In this respect, a study of the use of Olympus will in particular be included for the initial phase of RACE and the subsequent use of ESA experimental payloads for the implementation of IBCN.
- Allow for preliminary laboratory experimentation of system concepts (e.ge. modelling of antennas, simulations, software development).

The first two tasks would be done in liaison with or using input

from operating agencies and users in their field of interest and competence and deriving support from ad-hoc advisory groups of external experts.

II Development of payloads (DOP)

A first set of payloads is being identified through the General Studies and TPP programmes. Some preliminary work on elements of such payloads has been performed in the ASTP programme.

The candidate payloads so far under consideration for development and for possible flight opportunities in the 1991-1994 timeframe, identified as the "Technology Mission", are:

- Aeronautical/maritime mobile payload (ARAMIS) with phase of array multibeam for service demonstration and eventually possible operational use by e.g. INMARSAT
- Land mobile experimental payload (s) (with extension to aeronautical applications) for geostationary and other orbits (e.g. highly inclined elliptical)
- Optical inter-orbit links (Geo/Leo orbits: IOL) for future DRS
- Optical inter-satellite links (in Geo-orbit: ISL)
- Single access S-band payload with large antennas (5-8 m) for future DRS (development also applicable to land mobile)
- Multiple access S-band payload for DRS applications and possibly communication package(s)
- Propagation at 40/50 GHz and/or 80/90 GHz.
- Navigation and/or radiolocalisation experiment
- 20/30 GHz multibeam payload with reconfiguration and on-board processing
- Reconfigurable multinational and pan-European TV Broadcasting

This list is not exhaustive and other payloads might be identified in the course of the first year of the programme execution. Moreover not all missions mentioned in this list have reached the same level of technical maturity and/or acceptability by the potential users and consequently not all of them may be ready to fly on the 1991-94 "Technology Mission". Some payloads, although developed within PSDE programme, could be more conveniently flown on operational missions. Finally it is also possible that a specific payload, studied and developed within another programme is flown on this "Technology Mission", provided it is compatible with the timescale of the programme.

The final selection of payloads for this Technology Mission will depend on criteria such as:

- usefulness of the demonstration in-orbit of a particular technology or system concept
- urgency of market forces
- technical maturity of the payloads and the related hardware
- the level of consensus and acceptability of the experimental missions by the potential users of the experiments such as other space programmes (in-orbit infrastructure) and operating agencies, PTTs, etc...
- compatibility of the payloads from the viewpoint of development schedule, orbital position, frequency coordination, mass/power/volume availability on specific candidate platforms, etc.
- interest of the Member States, industrial policy and cost.

Those payloads, that are not selected for one or another reason for the first flight opportunities in 1991-94, but which are considered nevrtheless essential for achieving the objectives of the programme, will be developed either for a subsequent flight opportunity (AOTS), or for simulation testing on the ground, or to be directly incorporated in an operational mission at a later stage.

III Flight configuration studies and in orbit experimentation (CIE)

This programme line will start from the very beginning of the PSDE programme and will constitute its main backbone. The aim is the definition and realisation of an in orbit demonstration mission to be flown in the 1991-94 time period referred to as the "Technology Mission". The activities will include:

a) Identification of a set of suitable and compatible payloads for in orbit experimentation and of the candidate flight opportunities.

b) Comparative studies of specific space segment configurations.

c) Detailed Phase B definition studies, including phase B 1 (mission model definition) and phase B 2 (detailed configuration).

d) Development, launch and in-orbit operations of the selected experimental satellites and their payloads.

The first three activities would be executed in slice 2 of the programme. With regard to the fourth activity, a major decision, to be taken in 1988, concerns the authorisation to start developments towards the flight (classical phase C/D work) for the first satellite(s) of the Technology Mission. In parallel, a decision would also be taken on the continuation of the development of those payloads which would not be included in this first part of the Technology Mission, but which could then possibly fly on other satellite(s) of the Technology Mission or on AOTS after 1995.

A number of options for the payload/satellite configuration are studied in slice 2 of the programme. The most likely scenario, at present, consists of 2 (or 3) satellites in geostationary orbit and possibly a third (fourth) satellite in a highly inclined elliptical orbit. The allocation of payloads on the geostationary satellites will depend on different factors, such as ideal orbit position, frequency and mission compatibility, best use of resources etc.

It is envisaged at present that the configurations of the geostationary satellites will be as follows:

- SAT-1A using an Olympus platform and including:
  - a multi-channel TVBS payload for a television broadcasting agency (e.g. the Italian RAI)
  - an optical intersatellite link package for joint experiments with operating agencies (e.g. Intelsat and Eutelsat)
  - a radio-localisation payload (e.g. LOCSTAR).
- SAT-1B using a half ARIANE IV class platform and having ARAMIS as its main payload. Considering the interest of INMARSAT in a possible operational use of ARAMIS

payload in the transition phase between the second and third generation of INMARSAT system, other possible payloads which could be associated with ARAMIS should not make use of such advanced technologies that could delay the placing in orbit of this satellite.

SAT-1A and/or SAT-1B will nominally be placed in orbit in 1991.

- SAT-2, which should be placed in orbit in 1992, having a number of high technology payloads including:

  - an optical inter-orbit and inter-satellite link package as the main payload.

  - a single and multiple access S-band payload for future DRS applications.

  - possibly a land-seronautical mobile payload.

Studies on ways of allocating minor payloads, such as propagation and millimeter wave communication packages, to the two geostationary platforms, SAT-1B and SAT-2, will also be made with a view to making the total payload mass on each of the 2 satellites comparable. This would allow the use of identical platforms for the 2 satellites and the spare.

While SAT-2 would be completely financed by the ESA budget, SAT-1A and SAT-1B would rely on a shared cost arrangement with operating agencies (e.g. RAI, LOCSTAR, INMARSAT).

The third (fourth) satellite, in a highly inclined elliptical orbit, would be launched in 1993-1994. The possibility to use an Ariane V experimental launch in a multiple launch configuration will be explored for this demonstration mission. An accelerated programme of Advanced Research on Communication using Highy Inclined orbits for Mobile and other application (e.g. navigation) aiming towards Experiments And Demonstrations with the third Experimental Satellite of the Technology Mission (ARCHMEDES) would be actively performed in Slices 1 and 2.

IV Experiments and demonstrations with existing satellites

It is a programme line where the interest of Member States could be varying. It includes future activities of the same type as in APOLLO and PROSAT and should also support the RACE demonstration phase. The existing satellites which are considered for a first phase of this programme line would be the ones which are already or will be in orbit in the near future such as Inmarsat, ECS, Eutelsat 2, Olympus and national satellites.

A second phase would include the experiments and demonstrations with the satellites which will be developed within the PSDE programme. The preparation for these last experiments would, however, need an early start, in particular for those in orbit demonstration which might require preliminary on-ground experiments (e.g. optical interorbit and inter-satellite links.

Identified activities to be included in this line would be:

- Pilot experimentation in the data dissemination field. This activity could be a follow-on of the present APOLLO project, as well as an application of systems making use of microterminals.

- Development of new low cost earth terminals required for the experiment of future systems.

- Double-hop experiments of on-board processing (OBP) using an existing ECS satellite or Olympus or national satellites.

- Follow-on of PROSAT (to be defined).

- Other experiments, with Olympus or other satellites, for instance in support to the RACE demonstration phase, with associated developments of ground stations at 12/14 GHz and at 20/30 GHz.

- Development of flexible terminals able to operate with TDMA and future OBP systems.

- Preparation of the demonstrations to be performed in the second phase using PSDE satellites, in particular for DRS experimental and preoperational activities.

V Improvement of competitiveness

This programme line will include a number of activities aiming at the improvement of competitiveness of European products such as:

1) Improvement of existing platform designs. This activity should also include optimisation of satellite configurations for Ariane 5.

2) Support to export endeavours, through specific system and subsystem developments.

3) Development of techniques to reduce the cost of satellite systems

3.2. Advanced System and Technology Programme

A continuation of the ASTP is foreseen with programme slices renewable very 4 years. This activity will constitute the in-depth hardware counterpart of the prospective system studies of the Basic Support Line of the PSDE programme. This programme should contain competitive and non-competitive activities, with a percentage increase of the former.

It would cover new satellite system concept, platform and payload hardware as well as associated earth segment developmants. The activities of this programme line would mainly concentrate on initial development of breadboard models at unit or sub-assembly level.

Funding for this programme is assumed to continue at an annual level of 25 MAU.

3.3. Space Communication Infrastructure (DRS)

Those missions are, at present, identified as data relay services. They will, from the outset, be operational missions, in the sense that they will have to provide a reliable service. A European Data Relay Satellite (DRS) system is a necessary element of the IOI to ensure European independence in this vital area. A DRS System improves the return of the investment of low earth orbiting systems such as the Space Station Elements by dramatically increasing the real time communications capability with the ground per orbit and enabling very quick data processing where quasi-temporal effects are being observed which need rapid response. The NASA Tracking and Data Relay Satellite System (TDRSS) has amply demonstrated the operational advantages for low earth orbit satellites and Shuttle Management.

The use of a DRS will also enable the efficiency of ESA's conventional tracking and data acquisition network to be greatly enhanced.

The basic objective of the overall Data Relay Satellite programme is to set up a cost-effective infrastructure, in space and on the ground, in support to future European space programmes, which will provide, in a time frame compatible with the needs of users, the following near-continuous services:

- Transfer of data to and from low earth orbiting spacecraft and from launchers to ground controllers;

- Transmission of video and voice between low earth orbiting spacecraft and their ground control stations;

- Provision of telemetry and telecommand links between ground controllers and spacecraft in orbit;

- The capability to carry out ranging operations for orbit and position determination of spacecraft and, possibly, launchers during ascent.

DRPP

At the end of 1985 the ESA Council decided to start a Preparatory Programme for a Data Relay Satellite System (DRPP) with the following objectives:

- Establish and Define DRS Space and Ground Segments

- Investigate Feasibility/Benefits of International Cooperation

- Determine/Initiate Required Technology Developments

- Obtain Technical Baseline/Cost/Schedule for Implementing the European DRS Programme.

The DRPP will be executed in the time period 1986-88.

DRS-1

The DRS-1 programme consists of the development of an operational Data Relay Satellite System. A two satellite DRS system, orbited at, for example, 44°W and 61°E respectively, would provide very adequate coverage for IOI and other European missions. These satellites should be put in orbit before 1995, in time to provide the service for the IOI elements.

A decision on a specific configuration for the DRS space segment will be made as a result of analyses and trade-off studies made as part of the Data Relay Preparatory Programme (DRPP). A "Reference System Concept" is formulated and will be continuously updated to serve as a basis for comparison with alternative configurations.

The reference system concept envisages a spacecraft of the Ariane/SPELDA class with a microwave payload operating in the space to space link both at S-band, compatible with TDRSS, and at Ka-band for the high data rate services. The design of the spacecraft will also bu such as to be able to carry, in its first flight configuration, a pre-operational optical communications payload, in addition to the operational microwave payload.

Furthermore, it will have sufficient design flexibility and growth capability to be able to incorporate an operational optical communications payload on later flight models.

The spacecraft will be designed to be compatible with any position in the geostationary orbital arc and to be able to track LEO spacecraft up to 1000 km altitude).

The system should be capable of direct dissemination of the signals, to multiple mission centres for final processing.

The funding required for the system, over the period 1988/95, has been estimated at 560 MAU. This cost includes the development and manufacture of two flight units (DRS-1A and DRS-1B) and one onground spare, the launch of DRS-1A and DRS-1B and the initial control of the two satellites in orbit. It does not include the cost of the users' communication networks and ground user terminals.

Within the DRS programme a single mobile in-orbit user terminal for a Columbus element will be developed and qualified. This terminal will operate with both the space link frequencies of DRS-1 and would also allow interoperability with TDRSS.

The continuation of the operational mission will imply further space segment procurement in the course of the 90's. It is foreseen that a second generation of the Data Relay System, DRS-2, will be introduced before the end of the 90's using the technology and system concepts developed in the PSDE programme and therefore reaching an operational optical package for high capacity communications, longer antennas, multiple access techniques in S-band and maybe processing repeaters.

The design of DRS-2 will be such to ensure full compatibility with the user terminal operating with DRS-1.

3.4 Navigation Satellite System

The Executive's objective in this field is to foster cooperation with the user communities and the relevant international organisations and, if successful, to propose the setting-up, by ESA in cooperation with third parties, of a demonstration/preoperational system consisting of a certain number of satellites and earth stations. For the time being however, in view of the technical and institutional uncertainties associated with this proposal, the NAVSAT programme costs have not been incorporated in the proposed long term telecommunications programme.

3.5 Other Considerations

3.5.1 Cooperation

Cooperation with the USA and Japan has not been fully analysed in the scenarios described above. It is indeed difficult to envisage general cooperation in commercial communications given the competitive nature of this discipline. However, three specific areas of cooperation could be envisaged:

i) Data Relay, where either the European DRSS could be integrated into a world-wide system including Japanese and American satellites or at least a certain degree of interoperability of the different systems could be ensured

ii) Experimentation of optical ISL between an ESA and a NASA satellite; (maybe with ACTS modified to this end) or in cooperation with INTELSAT and EUTELSAT or others (e.g. Japanese operating agencies.

iii) Setting-up and operation of an experimental/preoperational NAVSAT system.

3.5.2 Interconnections among the Telecommunication Programmes

The PSDE programme, which involves essentially advanced technology items, is strongly connected to the evolution of DRS starting from its first configuration in 1994-1995. Three among the most critical possible elements of future Data Relay Satellites, i.e. a single access payload including a large S-band antenna, a multiple access S-band payload, and the optical payload for In-Orbit Communications between the DRS and the user satellites and platforms in LEO, are considered for in-orbit experimentation in the frame of the PSDE Programme.

The ASTP programme is also strongly connected to the PSDE programme. The prospective technological work for payload development was and will be carried out in ASTP, particularly for those payloads that will not be embarked on the first flight opportunity. The ASTP will also contribute to the preparation of AOTS.

4. FUNDING

The funding of the different programme areas is summarised below. The corresponding expenditure profile is given in Table 1, the schedule planning is shown in Figure 1.

5 SUMMARY (time period 1987-1995)

o PSDE	997 MAU
- I Basic Support Line (BSL)	
- II Pre-development/Development of Payloads (DOP)	
- III Flight Configuration Studies and In-orbit Experimentation (CIE)	
- IV Experiments and demonstration with existing satellites (CIE)	
- V Improvement of Competitiveness	
it includes in-orbit demonstration missions :	
- Technology mission (3/4 satellites) 1991-94	
- Initial development of AOTS (2 satellites) 1997-98	
o DRS	
- DRPP	29 MAU
- DRS-1 operational mission (2 satellites by the mid-90s)	560 MAU
- DRS Leo Terminal	55 MAU
o NAVSAT	
Decision will depend on the success of the setting up of an international institutional framework	

Table 1
Financial Planning
Profile of expenditure for the telecommunications programmes 1987 — 2000 (MAU in 1985 e.c.)

PROGRAMMES	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	TOTAL
Telecom 3 bis	55 2	31 2	22 9	16 9	14 9	11 9	6 0	6 0							165.0
Olyapus Phase C D	85 2	62 2	4 7	4 6	4 3	4 3									165.3
ASTP 2	18 2	5 8													24 0
PROSAT 2	3 2	1 4													4 6
Apollo	3 0														3 0
TPP	2 9														2 9
IOC	7 1	4 9													12 0
ASTP 3	16 4	23 5	25 3	22 3	19 6										107 1
DRPP	10 0	13 0	6 0												29 0
MARECS B B2	2 0	1 7													3 7
PSDE Slice 1	3 9	5 9	5 9	4 9											20 6
PSDE Slice 2	21 6	54 0	28 5												104 1
TOTAL APPROVED TELECOMN	228 7	203 6	93 3	48 7	38 8	16 2	6 0	6 0							641 3
ASTP 4					5 0	20 0	25 0	25 0	25 0	25 0	25 0	25 0	25 0	25 0	225 0
PSDE 3 8	3 0	84 0	110 0	135 0	155 0	140 0	120.0	125 0	125 0	165 0	185 0	185 0	135 0	135 0	1802.0
DRS 1			30 0	53 0	67 0	90 0	110.0	105 0	105 0	35 0	15 0				610.0
DRS OPS									5 0	10 0	10 0	10 0	10 0	10 0	55 0
DRS 2													50 0	90 0	140 0
DRS LEO Terminal			5 0	10 0	12 0	15 9	7 0	5 0							54 9
TOTAL FUTURE TELECOMN	3 0	84 0	145 0	198 0	239 0	265 9	262 0	260 0	260 0	235 0	235 0	220 0	220 0	260 0	2886 9
GRAND TOTAL	231 7	287 6	238 3	246 7	277 8	282 1	268 0	266 0	260 0	235 0	235 0	220 0	220 0	260 0	3528 2

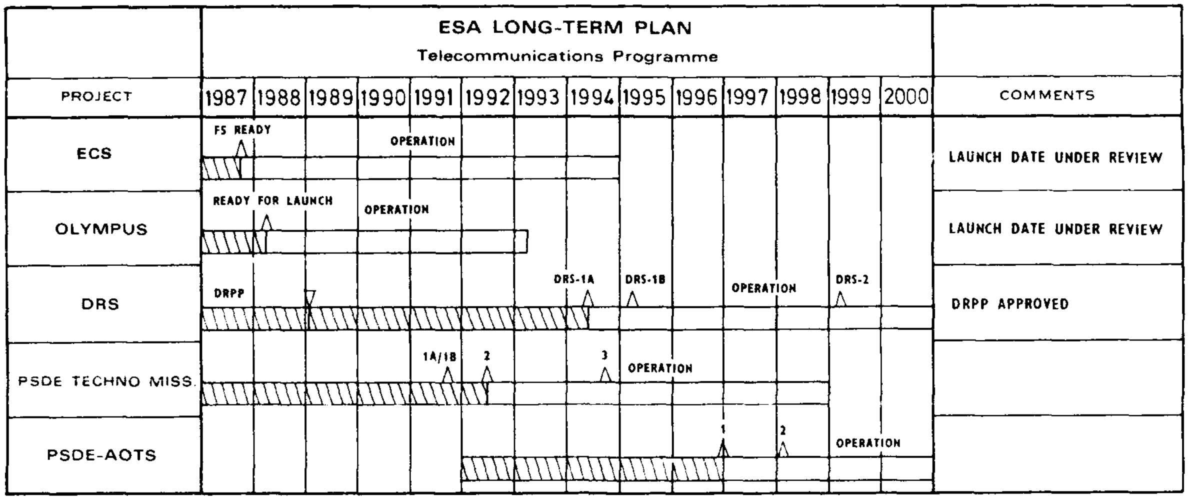
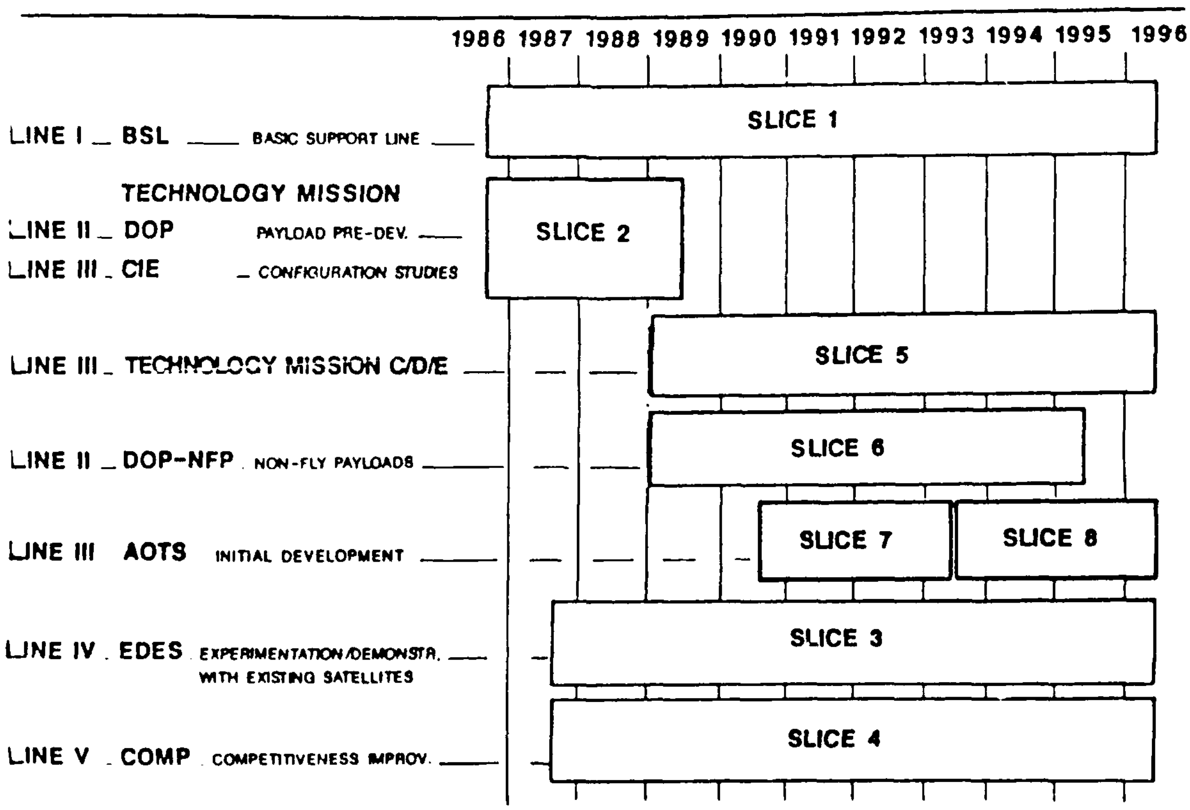


Fig.3 Time Plan of the Telecommunication Programmes



PSDE: Programme lines and slices

CHAPTER 10

POSSIBLE SATCOM SYSTEM ARCHITECTURES

10.1 DESIRED CHARACTERISTICS

In order to limit the range of SATCOM architectures to be developed and discussed with a view to bringing out the most promising candidate architectures for NATO it was agreed that some indisputable/generic requirements for a future NATO SATCOM system had to be defined as follows:

(i) Coverage: NATO areas including the polar region, in which terminals of various size and traffic capacity will be used including terminals for submarines.

(ii) Orbits: consistent with the coverage to be provided and minimum cost.

(iii) Transmission delay: consistent with the transmission requirements of the overall system.

(iv) Launch: capability to use a wide range of launch vehicles and sites consistent with availability, low vulnerability and cost.

(v) ECCM: have a high degree of AJ capability flexible enough to exchange traffic capacity with AJ performance and interoperability. Low Probability of Intercept (LPI) is also a requirement, particularly for mobile users (higher frequency bands, lower transmitter power, band spreading).

(vi) Resistance to physical attack: cost effective mixture of some or all of the following: hardening, proliferation, manoeuvre capability, active defence.

(vii) Adaptability: capable of evolutionary adaptation to changes in NATO/national requirements (including operating orbits) and threat.

(viii) Transition: architectures allowing a smooth transition from existing to future systems are preferred (taking into account the investment in the ground terminals and control stations).

(ix) Availability: ideally continuous service is required and this can be approached by having internal redundancy in spacecraft, in orbit spare spacecraft, repair of spacecraft and refuelling and the use of friendly national or international spacecraft systems.

(x) Interference: an architecture which minimizes interference and hence coordination difficulties will be preferred.

(xi) System autonomy: architectures which require the minimum of ground control are preferred. Autonomous control with manual fallback is desirable for all routine, failure and other selected modes to maximize service availability and minimize control costs.

The two main interacting ingredients of a SATCOM system architecture to serve a given ground environment are the orbits and the characteristics of the satellites. The orbits determine the coverage and the number of satellites to be used for a given service continuity and availability and the satellites, either singly or combined, determine the capacity, reliability and survivability of the system.

From a perusal of the literature and discussions held within the group the following system architectures were identified for initial consideration and evaluation by the group:

(i) Systems using satellites in inclined elliptical orbits:

(a) Molniya

(b) Tundra

(c) Loopus

(ii) Satellite Cluster Systems

(a) Cloudsat

(b) MEWS

(iii) Lighsat and Proliferated (Multiple) Satellite Systems (MSS) in LEO.

(iv) Geostationary satellite systems and the use of tethers.

(v) Systems of satellites in combinations of orbit types

(a) Geostationary plus circular polar

(b) Geostationary plus 24-hour Molniya

These systems will be discussed and evaluated below taking into account the perceived NATO requirements postulated above.

10.2 INCLINED ELLIPTICAL ORBITS (Molniya, Tundra and Loopus types)

10.2.1 Introduction

The NATO SATCOM II, III and IV systems have all used geosynchronous circular orbits with a small inclination to the equatorial plane. Ideally, the orbits would have been truly geostationary, i.e. with zero inclination, but it was found more cost-effective to tolerate a change in inclination of around 7° over the spacecraft's lifetime than to accomodate the additional fuel needed for North-South station keeping. The attraction of a quasi-geostationary orbit is that the satellite always appears at more or less the same position in the sky. Hence the ground stations do not have a severe tracking requirement and, moreover, a single satellite can provide continous service to all ground stations within its "footprint" covering about one-third of the Earth's surface. For NATO, this means that such a satellite positioned at around 18° W longitude can provide continuous coverage of most of the area of NATO operations. It is visible at useful elevation (at least 8°) from all 21 of the large static ground terminals of the current NATO SATCOM system.

Clearly the case for continuing to use a single active satellite in geostationary or quasi-geostationary orbit in the post-NATO IV era is strong. However, there are several reasons why alternative orbits are at least worth considering, in particular:

(i) By the year 2000, it will no longer be possible to discount the possibility of physical attack on a satellite in geostationary orbit.

(ii) New operational requirements may imply a need for the extension of SATCOM coverage, particularly to polar latitudes.

(iii) The problems of coordinating satellite locations on the geostationary arc and of avoiding interference are becoming severe, particularly at SHF.

(iv) At EHF, the low elevation at which a geostationary NATO satellite is seen from North America and Northern Europe means that SGTs in these locations will suffer high atmospheric attenuation and be particularly susceptiple to outages in rain.

(v) As seen from a geostationary NATO satellite, the angular separation between European SGTs and Soviet-bloc jammers is very small. Thus spaceborne nulling antennas need very high resolution (and a very stable platform) to be effective.

Systems of satellites in inclined elliptical orbits can provide service within a coverage region that is better matched to NATO's requirements than that of a geostationary satellite, while at the same time avoiding the congested geostationary arc. Survivability is enhanced, at least in the sense that an enemy must now destroy more than one satellite in order to deny service completely.

An analysis in terms of numbers of satellites and types of different orbits for providing continuous coverage of the region of interest to NATO is presented in Appendix 10A from which the following can be deduced:

10.2.2 Principle of the Inclined Elliptical Orbit

Inclined highly-elliptical orbits (HEOs) have been used by Soviet "Molniya" satellites for many years. At apogee, such a satellite is typically around geosynchronous altitude whereas at perigee its altitude may be only a few hundred kilometers. Near apogee the satellite's movement more or less keeps pace with the rotation of the earth for several hours and thus it maintains a fairly constant "footprint" during this period. Continuous coverage can therefore be provided using only a small number of active satellites.

A major advantage of the Molniya orbit is that the apogee boost motor used to circularise the orbit of a geosynchronous satellite is not required: typically this accounts for about half the mass of the satellite at launch.

Fig.10.1 shows how the position of a HEO satellite along its orbit, measured in polar coordinates from the centre of the earth, varies with time. Coordinate values for a circular orbit are shown for comparison. The relatively slow change in the position vector near apogee can be clearly seen in the case of the HEO. In this example the HEO has a period of 12 hours and a perigee altitude is 1000 km, corresponding to an eccentricity of 0.726.

In a system with 12-hour orbits, each satellite provides coverage of the service area at alternate apogees, i.e. every 24 hours. Because the operational period of each satellite is necessarily less than the orbital period, at least three satellites are needed to provide continuity.

The ground trace of a satellite in a 12-hour highly-elliptical orbit inclined at an arbitrary angle of 52° is shown in Fig. 10.2. The first apogee is at 22° W. Points on the ground trace indicate the time in hours since the initial perigee. It can be seen that for at least four hours either side of the first apogee the satellite is well placed for coverage of the NATO area. A system with three active satellites in such orbits could therefore provide continuous service to NATO.

Systems with orbital periods greater than 12 hours are also possible. The basic requirements are that all satellites in the system should follow the same ground trace, and that there should be sufficient satellites for at least one to be visible at all times from the desired service area. This could be achieved, for instance, with two satellites in 24-hour orbits. Consideration has also been given to a system based on 16-hour orbits in which each satellite provides service for eight hours around the relevant apogee. Since it takes 48 hours for a satellite in a 16-hour orbit to repeat its ground trace, a total of six satellites is required in this case.

10.2.3 The Implications of the Choice of Inclinations

Because the Earth is not a perfect sphere, each of the classical parameters defining a satellite orbit changes with time. However, it can be shown [10.1] that whereas the semi-major axis a, the inclination i and the eccentricity e have the same values at the beginning and end of each orbital period, the Right Ascension of the Ascending Node and the Argument of Perigee in general do not. Thus there are two major effects, a tendency for the orbital plane to precess about the Earth's axis and a tendency for the ellipse to rotate within the orbital plane. In the systems of interest which require the orbit to be synchronised in some sense with the rotation of the Earth, the first effect can be compensated for by a small adjustment of the orbital plane. However , the second effect will require correction, for example by periodic thruster firings, if the coverage and availability of the system are to be maintaiend.

It is shown in [10.1] that the change in the argument of perigee, ω, in each orbital period is given by:

$$\Delta\omega = 3\pi J_2 (R_e/p)^2 (5\cos^2 i - 1)/2 \qquad (1)$$

where p is obtained from the relation $p = a(1-e^2)$, Re is the mean radius of the Earth and J_2 is a constant equal to 1.082637×10^{-3}. It is evident that $\Delta\omega$ is zero when $\cos^2 i = 1/5$, i.e. when $i = 63.43°$.

Thus if the orbit is chosen to have this critical inclination the rotational tendency is eliminated, at least to a first order. Soviet Molniya satellites use this inclination.

For other inclinations, it is possible to calculate the change in velocity Δv needed at the appropriate point on each orbit in order to restore the riginal value of $\Delta\omega$, and hence to determine the amount of fuel needed for the corrective manoeuvre. For small $\Delta\omega$, it is shown in Appendix 10B and in [10.2] that:

$$\Delta v = 2.a.n \frac{e}{\sqrt{(1-e^2)}} \sin (\Delta\omega/2) \qquad (2)$$

where $n = 2.\pi/P$ and P is the orbital period.

The amount of fuel required to produce a given change in spacecraft velocity is determined by the "specific impulse" of the fuel. This is defined as the length of time that 1 kg of fuel can produce 1 kg of thrust. If the spacecraft mass is M, the mass m_f of a fuel with specific impulse I required to produce a change in velocity Δv is therefore:

$$m_f = M \Delta v / (I.g) \qquad (3)$$

where g is the acceleration due to gravity.

For Hydrazine used as a monopropellant (as in the NATO IV spacecraft) the specific impulse is 230 seconds

Expressions (1), (2) and (3) have been used to calculate the mass of Hydrazine needed per day to maintain spacecraft in a variety of inclined highly-elliptical orbits. This mass is found to be unacceptably large even for small deviations from the critical inclination. For example, a 1000 kg spacecraft in a 12-hour orbit with an inclination only 0.3° greater than the critical angle would need to use 0.4 kg of fuel per day to overcome the rotation. For a 12-hour polar elliptical orbit the requirement increases to a massive 4.6 kg/day.

It is concluded that a future NATO SATCOM system based on inclined orbits would need to use the critical inclination of 63.4° unless some form of propulsion could be employed that was much more efficient than those of today. Use of other inclinations would only be realistic for short-duration missions where the orbit could be allowed to rotate without correction, for example if the system was deployed "on-demand" to provide short-term emergency communications in war.

Even though one may achieve great economy in weight and maintain a precise orbit by using "ion thrusters", it can be stated that 63.4° orbit can be the type preferred for NATO and that it does not have any disadvantages at all

10.2.4 Possibilities for a "No-Break" Service

A fundamental problem with a system of two or more satellites in inclined elliptical orbits is the need for the ground segment to cope with a change of operational satellite several times per day. Terminals with essentially tactical roles can arrange to change satellites in between message transmissions, but those supporting multi-channel common-user bearers may need to provide a "no-break" service. One possibilty is to provide each SGT with two RF heads so that both satellites can be tracked simultaneously prior to switch-over, but in the case of the NATO large static SGTs this would prove extremely expensive.

An alternative possibility is to arrange for the rising satellite (that coming into service) and the setting satellite (that going out of service) to pass sufficiently close to each other that, briefly, both appear simultaneously within the beam of the SGT's antenna. At the moment of closest approach, the transponder of one satellite could be turned off and that of the other satellite turned on. At the same time the SGT would cease tracking the setting satellite and begin to track the rising satellite. Provided the system could tolerate, or compensate for, the discontinuities in delay and Doppler shift at the moment of switch-over it would be possible to maintain links through the transition without interruption.

In the 12-hour elliptical orbit considered in Fig. 10.2, the ground trace crosses over itself some 15 minutes before and after perigee. It is found that orbits of this type have 0, 1 or 2 cross-over points depending on the inclination of the orbit, and that the timing of the cross-overs depends strongly on the inclination. By careful choice and maintenance of orbital inclination it is possible in principle to make a cross-over occur exactly 4 hours either side of apogee, so that the rising and setting satellites pass simultaneously through the same point in space at that time.

The inclination α required for cross-over at a specified time can be found by requiring the longitude lg_s of the sub-satellite point to be the same τ hours before (or after) apogee as it is at apogee. It can be shown that:

$$\sin \alpha = \frac{(\cos^2 \nu(\tau) - \sin^2 \xi(\tau))^{1/2}}{-\cos \nu(\tau) \cos \xi(\tau)} \qquad (4)$$

where $\nu(\tau$ is the angular position of the satellite measured from perigee, obtained from consideration of Kepler's laws (see Appendix A of [10.2] for details), and

$$\xi(\tau) = \pi(1 - \tau / 6)/2$$

The required inclination as given by equation (4) is shown in Fig. 10.3. as a function of τ for the case of a 12-hour orbit with perigee altitude of 1000 km. It can be seen that for inclinations less than about 62° the cross-over occurs 30-40 minutes after perigee. For inclinations between 62° and 65° there are two cross-overs, and for inclinations greater than 65° no cross-over occurs at all. For a cross-over 4 hours either side of apogee an inclination of 64.73° is required. It is interesting to note that this inclination is close to the critical value for orbit stability derived in section 10.2.3.

The inclination required for a conjunction four hours from apogee is also dependent on perigee height. Fig. 10.3. includes a curve of inclination versus τ for a 100 km perigee altitude as well as the 1000 km case just discussed. In the case of a 12-hour orbit it would be possible to choose the perigee altitude so that a conjunction 4 hours from apogee corresponded to the critical inclination of 63.43°. The required altitude is found to be 1330 km. A ground trace for this case is included in Fig. 10.4: It cat be seen that during the 8-hour operational period the ground trace forms a narrow loop covering about 20° of latitude.

The switch-over from one satellite to another must take place during the time when both satellites are simultaneously within the main beam of each SGT. To determine the duration of this "window", the separation of the satellites when close to conjunction can be calculated as a function of time and compared with the antennas beamwidth. If the two satellites are seen at azimuth angles Az_1 and Az_2 and elevation angles El_1 and El_2 then the angular separation ϕ is given by:

$$\cos \phi = \cos El_1 \cos El_2 \cos(Az_1 - Az_2) + \sin El_1 \sin El_2 \qquad (5)$$

The elevation of each satellite as seen by an observer at longitude lg_p and latitude lt_p can be shown to be:

$$El = \tan^{-1}((r_s \cos \theta - re) / (r_s \sin \theta)) \qquad (6)$$

where:

$$\cos \theta = \cos lt_s \cos lt_p \cos(lg_p - lg_s) + \sin lt_s \sin lt_p \qquad (7)$$

The azimuth is found from:

$$\tan Az = \frac{\cos lt_s \sin(lg_s - lg_p)}{(\sin lt_p \cos lt_s \cos(lg_s - lg_p) - \cos lt_p \sin lt_s)} \qquad (8)$$

In these expressions r_s is the range of the satellite measured from the centre of the earth, r_e is the radius of the earth and lt_s and lg_s are the latitude and longitude repectively of the sub-satellite point. r_s, lt_s and lg_s may be determined as functions of time from considerations of the dynamics of the elliptical orbit and its parameters.

Expression (5) has been evaluated in the vicinity of cross-over for a typical system based on 12-hour highly-elliptical orbits. The observation point was taken as Kester (Belgium) but the result is not expected to depend strongly on this. It is found that the angular separation changes at a rate of 0.54° per minute. Thus for a 12 m SGT with a beamwidth of 0.22° the two satellites will only be within the beam together for 24 seconds. In the case of a 2.4 m SGT with a beamwidth of 1.1° this increases to approximately two minutes.

The existence of even this brief period is dependent on precise maintenance of orbital parameters, in particular the inclination. If the inclination is slightly in error, the ground trace will have a cross-over point somewhat more or less than 4 hours from apogee, but because the satellites are spaced eight hours apart there will no longer be a conjunction at this point. Minimum separation will still occur four hours from apogee but unless that separation is less than about half a beamwidth service will be interrupted while the SGT acquires the new satellite. The shift in cross-over timing for a given inclination error can be found from Fig. 10.3. If this shift is t_s minutes then (provided the error is small) the minimum separation can be estimated by determining (from eqn. 5) the angle between the position of the rising satellite t_s minutes before cross-over and that of the setting satellite t_s minutes after cross-over. This has been done for a typical 12-hour orbit and it is found that the minimum separation is 1.3 times the inclination error. Thus an error of only 0.08° is enough to reduce the duration of the switch-over period to zero in the

case of a 12 m SGT.

Alternatives to the 3-satellite, 12-hour HEO system have been examined to see if they could provide longer "windows" of simultaneous accessibility during conjunctions and/or reduced sensitivity to inclination errors. In general these alternative systems suffer from the fundamental disadvantage that the "periodic conjunction" requirement can no longer be satified with a nominal inclination of 64.43°. It is found that while the "window" can be extended somewhat the sensitivity to inclination errors is no better than that of the 12-hour system. For example, periodic conjunctions could be realised every 8 hours in a six-satellite system based on 16-hour orbits inclined at 77.5°. In this case the angular separation changes at a rate of 0.21° per minute near the cross-over point, giving a "window" of 63 seconds for 12 m SGTs and about 5 minutes for 2.4 m SGTs. The sensitivity of the minimum angular separation to inclination errors is 1.4, similar to that for the 12-hour system.

One consequence of the precise control of orbit parameters discussed above could be to make the probability of a collision significant. This would need careful consideration.

10.2.5 Potential of Inclined Elliptical Orbits for NATO

The extent of coverage of the three-satellite 12-hour HEO system with 63.43° inclination and with the eccentricity chosen to cause conjunctions every 8 hours is shown in Fig. 10.5. The apogee longitude is 22° W. The coverage area is defined as that from which the operational satellite is always visible at at least 10° elevation. It can be seen that the coverage area includes almost all of the territory of the NATO nations as well as the North Atlantic, Mediterranean and North Polar regions. It is clear that the coverage is better matched to NATO's requirements than that of a geostationary satellite, which is shown in Fig. 10.5. for comparison. Furthermore the operational satellite will in general be seen at much greater elevation from the existing static NATO SGT sites than will a geostationary satellite.

While not able to satifsfy the "periodic conjunction" criterion, elliptical 24-hour orbits inclined at 63.4° may be still useful for tactical roles, particularly as they offer good coverage of both the NATO and polar regions with a system of only two active satellites. Fig. 10.4. includes the ground trace of a satellite in a 24-hour orbit with the critical inclination and a perigee altitude of 1000 km ("24-hour Molniya"). Apogee is at 22° W longitude. It can seen that the trace is fairly linear and covers some 110° of longitude during the 12 hours that the satellite would be operational in a two-satellite system. This compares with the compact loop traced by satellites in 12-hour Molniya orbits, also shown in Fig. 10.4. The result is a reduction in the width of the system coverage area, as shown in Fig. 10.5. Nevertheless, the coverage is still extremely useful for NATO, including all the territory of the NATO nations except the south-western USA, and the entire north polar region.

The coverage of the 63.4° inclined, 24-hour orbit system may be improved by varying the perigee altitude (and hence the eccentricity) in an attempt to make the ground trace more compact. This is the principle of the "Tundra" orbit which is being considered for a Canadian military SATCOM system [10.3]. In the example considered here, the perigee altitude has been increased to 20000 km. From Fig. 10.4. it can be seen that the ground trace now crosses over itself, and is restricted to an area measuring about 45° in longitude by 30° in latitude during the 12-hour operating period. As shown in Fig. 10.5, the resulting coverage area is significantly extended to the south, east and west relative to the "24-hour Molniya" system, but more restricted in the north due the increased latitude excursion. This loss of northern coverage offsets the advantage to NATO if the apogee longitude is held at 22° W, but is evident from the figure that if the apogee were moved to around 38° W almost all the territory of the NATO nations would be covered. Appendix 10C contains a reproduction of [10.4] where Tundra orbits are considered for a possible NATO application.

It is interesting to note that a system called LOOPUS which stands for "Geostationary Loops in Orbit Occupied Permanently by Unstationary Satellites" is being studied by the Deutsch Bundespost (German PTT) for a public mobile satellite network covering the northern hemisphere. Consideration is being given to the use of 12-hr and 16-hr satellites in 63.4° orbits. This system is outlined in Section 9.6 and in Appendix 9B.

10.3 SATELLITE CLUSTER SYSTEMS (CLOUDSAT and MEWS)

10.3.1 Introduction

Satellite clusters are systems of satellites in which the application functions of a single large satellite are distributed spatially among a series of small satellites which operate in the same or adjacent orbits. For the configuration to remain in a stable spatial inter-relationship, the sub satellites have their individual positions controlled by a nominated master satellite in the cluster using a communications link between the attitude control system of the master and those in each satellite of the cluster. The link may also be used to provide other functions such as extended redundancy of for example telemetry, command and control if there is a loss of these functions in any individual satellite.

A variation of the basic concept is a system of satellites in which each sub-satellite of the cluster has an identical function but is only operational for a fraction of time. In this instance the application of the cluster can be both time and spatially distributed.

A combination of the two concepts is possible and a multi-role function e.g. communications and surveillance could also be undertaken. A diagrammatic representation of a Cluster system is shown in Figure 10.6.

The attributes of an ideal military SATCOM system in addition to meeting it's communications requirement may be summarised as follows:-

(a) Have minimum development, recurring and launch costs.

(b) Be capable of up grade and expansion whenever required to meet new or different traffic demands.

(c) Permit defective and life expired elements of the system to be replaced in orbit without man intervention.

(d) The refuelling of satellite elements when ever necessary to extend the life time of a still functioning system, again without in orbit intervention by man.

(e) Have virtually zero down time at low cost.

(f) Obviate field of view problems and minimise interference in the frequency spectrum.

(g) Make maximum use of orbital slot allocations.

(h) Have a low vulnerability to physical attack

(i) Have enhanced anti-jamming performance.

(j) Be useable in either geostationary, low earth or eccentric orbits.

A well designed cluster system of satellites can meet these requirements in large measure.

10.3.2 The MEWS System.

A system of satellites known as Marconi Earth Watch System (MEWS) devised for Earth Observation from either GEO or LEO polar orbits will illustrate the concept of a cluster system which may serve as a model for a Military SATCOM system. Figure 10.7 shows the basic configuration possibilities for this system.

10.3.2.1 The System Concept.

The concept is based on the principle of spatial distribution. The basic element is a satellite and the system may consist of one or several satellites which may operate either autonomously or synergistically in any number of sets of groups which are required operationally. Each element of the distributed system is connected through a network so as to maintain the fullest operational services by optimum use of the resources of all the elements. The networking can be implemented by a variety of alternative means which are dependant on the separation of the elements in the cluster and by the type of resource to be exchanged between satellites.

These are shown in Figure 10.7 There are three basic building bricks of the system.

a) Satellite Platform or Bus.

The platform is conventional in so far as it supplies the utilities of power, telemetry command and control (TT&C), attitude and orbit control (AOCS) with redundant subsystems its satellite element. Normally one platform of the cluster will have the role of master controller of the complete cluster system. Each platform has an identical AOC system and can therefore undertake the role of master in the event of failure. Control signals for this purpose are fed to and from each satellite over the link network.

The platform is constructed on a modular basis at all levels. The modularity is such that the needs of each platform and its payload may be closely matched from the smallest number of modules. For example solar arrays are constructed from a number of panels, but the number of panels used on a particular spacecraft would be those required to meet the needs of that specific satellite.

b) Payloads.

Payloads for each satellite in the system may be identical and/or different. This will depend on the nature of the overall system. Whenever they can be identical, the payload construction would be modular so that redundant payload elements could be shared within the overall group if needed. If the payloads were different in different satellites as might be the case for different frequency band use, effective redundancy would be available by inter-satellite sharing with in-orbit spare satellites. For each cluster system an optimum redundancy implementation would exist to satisfy the overall operational requirement. Figure 10.8 illustrates a cluster configuration and the available redundancy pathways which are additional to the internal redundancies within each spacecraft.

c) Satellite Terminal or Hub.

The Satellite Terminal consists of a satellite with a payload which serves as a distribution centre for inter-satellite signals. The terminal permits the sharing of satellite resources when necessary and is the orbit base for cluster refuelling.

The Satellite Terminal is the first element of the cluster system to be placed into orbit. The system is then progressively extended by placing further satellites in orbit and linking them to the terminal via the appropriate link. Identical docking/undocking ports on all satellites, including the terminal, enable spent spacecraft to be removed from the cluster and replacements re-integrated into the system. The spent spacecraft are either de-orbited earthwards to burn up in the atmosphere or ejected into deep space to avoid orbital clutter. At some future time it may be possible to recover spent spacecraft economically and after refurbishment, be re-used.

10.3.3 Network Links.

The links between elements of the system normally take place via the satellite terminal, and dependant on the nature of the link enable different functions and operations to be carried out by the cluster. If the satellite terminal is out of action or being replaced when life expired, the cluster network is re-configured to allow a nominated satellite to act as the signal distribution centre. All functions of the system except refuelling will continue during this time.

10.3.3.1 Hard or Rigid Links.

These are made by direct docking of the satellite with the terminal. A long hard link is formed by docking a telescopic probe with the Hub and subsequently extending the probe length. A short link is also formed by docking a satellite with the Hub. The short link is used for fuel transfer and to construct a larger single spacecraft built around the Hub. The short link would therefore carry signals, power and fuel for both payloads and platforms. The long link would serve similar purposes except for fuel transfer. It's main use would be to avoid mutual interferences between fields of view of antenna systems, and for thermal access to deep space without periodic shadowing by parts of the satellite.

10.3.3.2 Flexible or Soft Links.

These links are also made by docking satellites with the terminal and then reeling out flexible cable or optical fibre as the spacecraft separate and take up their stations. In the case of cable this would carry power, control and communication signals while only signals would be carried by fibre. Cable can also be used as a tether between spacecraft and could for example be used to position satellites above and below geo-stationary orbit along an extended earth radius, thus permitting more satellites to be positioned at the same geostationary location (see Section 10.4 on Tethered Systems). The link length might be of the order of 1 to 5 km.

Optical fibre would be used where a high degree of attitude and positional freedom was required as might be the case when avoiding physical attack, while at the same time maintaining signal or control links between cluster elements. The link length in this case might be between 1 and 10 km.

10.3.3.3 Microwave and Laser Links.

These links are made between two way terminals on both satellites and Hub, and would carry communications and control signals. Initially the links would be formed by a search and lock on procedure and maintained by subsequent tracking between elements. Laser links would be used in preference to microwave links if large signal bandwidth was needed and yet be virtually immune to jamming. There may however be a cost/mass advantage in using EHF inter-satellite links in the near term.

Both forms of link would be used to reconfigure the system in the event of Hub damage or failure. Links would then be made in a ring or star format directly between satellites, one of which would assume control of the cluster. Links of this type could also be used for inter-orbit communication between TUNDRA spacecraft and those in geostationary orbit thus providing another, but indirect route for polar communications under jamming conditions. Link lengths can be several hundreds of kilometres if required.

10.3.4 System Charcteristics

Many of the benefits which can be achieved from a cluster system arise from the inherent attributes of spatially distributed function and communication links between elements of the cluster.

Distributing the functions spatially leads automatically to more but smaller satellites. This in turn leads to a scale of modularity which allows the system to be altered in small increments and only in those areas which require change. As a result whenever up grade or repair is necessary, this can be implemented at the lowest cost in down time and money.

Networking the satellites by several means not only re-unifies the system to perform as a single entity, but enables each part to operate autonomously or as a number of co-operating entities.

Both attributes lead to a system which degrades gracefully and can regain its initial performance with near zero down time and in a timescale which matches likely launcher availability.

10.3.4.1 System Repair

Repair of satellite systems have so far been by substitution of an in orbit spare or by the launch of a new replacement spacecraft. Recently recovery and limited repair has been demonstrated as feasible at altitudes around 400 km by US astronauts operating from STS (Shuttle). Recovery and/or repair at higher orbits await the emergence of economic orbit transfer vehicles and sophisticated robotics. For Military SATCOM recovery and repair by manned means as in Shuttle, would require the descent of the sick spacecraft to a orbit similar to shuttle orbit to avoid the radiation hazards at geostationary or intermediate altitudes. The repaired craft would then be re-launched back into the operating orbit. It is unlikely that such a procedure would be more economic than the present methods.

The system of repair proposed for a SATCOM cluster follows the current practice of repair by satellite replacement, using simple robotic docking with the terminal to complete the inter-satellite links. Because of the reduced size and singular operational function of each cluster element, and the increased redundancy available via networking, a replacement satellite and its launch will be cheaper than for a single multi-function one.

The size of each satellite of the cluster is chosen to be capable of carriage by a range of launchers in either single or piggy-back configurations. This ensures the maximum choice of launcher, and hence the lowest available launch cost and the earliest available launch. The cluster system would normally have sufficient redundancy to avoid any part system down time while awaiting a launch of a replacement element. The in-orbit spare of a conventional system is essentially always available in distributed form in a cluster group.

10.3.4.2 System Up Grade and Growth.

System up grading is carried on the same basis as repair by removing an existing satellite and replacing it with a new satellite which has the enhancement features required. System growth is obtained by adding further satellites to the system until limited by the number of links available in the Hub. Further growth would then be implemented by adding a satellite terminal via a short rigid link to the existing terminal. Further satellites are then docked and linked into to the new terminal.

The size of the system is ultimately limited by the complexity of the management system required for station keeping, attitude control and the autonomous allocation of resources between elements. (See section 8.10 on Robotics). At that point additional clusters could be introduced. A new clusters could be connected to the first cluster by microwave or laser links if required but inter-operations would only take place at high level.

10.3.4.3 Re-fuelling.

Satellite life is frequently terminated by fuel exhaustion. The life of the other elements of a satellite are notoriously difficult to predict. The electronic systems of some SATCOM satellites have continued to function within specification for more than six times the design life time, but advantage of this extended life could not be taken as the fuel required for the attitude and orbit conrol functions was lacking, resulting in uncontrolled off station drift.

In such circumstances an ability to refuel a satellite would be an advantage and for a cluster system even more beneficial. Re-fuelling satellite elements of the cluster would be carried out by docking a tanker spacecraft with the terminal. Each cluster spacecraft would also dock with the Hub in turn. Fuel transfer from the tanker to the satellites would then take place via the docking ports. When the re-fuelling operations were complete the tanker spacecraft would be de-orbited. The tanker spacecraft would utilise tha same platform as other satellites of the cluster but because of its short operational life would have much reduced solar arrays and power systems and would be optimised to maximise the fuel tank capacity.

10.3.4.4 Reduced Vulnerability to Threat.

a) Physical Attack.

A cluster of spatially disributed satellites is equivalent to a set of multiple targets for any form of physical attack be it from A-Sat weapons, kinetic energy, beam or nuclear weapons. As such more enemy resources would be required to attack and destroy the system. The intersatellite links allow the transfer of communications traffic to the surviving satellites and thus extends the time for which critical SATCOM services remain available.

b) Anti-Jamming performance.

The spatial distribution and inter-satellite links of a cluster provide a method of increasing the enemy resources required to jam the system. In a jamming attack, communication signals are distributed pseudo randomly multiplexed in space and time division for both up and down link signals. Depending on the separation between satellites of the cluster, and all satellites of the cluster would require to be jammed as it is not possible to predict which satellite is in use until too late. Different pseudo multiplexing codes might be used for EHF and SHF.

In a variation (subset) of the MEWS concept called CLOUDSAT, use is made of several satellites which work in the receive mode and are spatially separated so that one jammer cannot jam more than one satellite at a time. These receiving satellites transmit via cross- links to a common downlink which may be supported by one or a few satellites. The system uses as many satellites as are necessary to gain the required jamming advantage. The use of multiplicity of satellites, as a bonus, also provide a degree of in-orbit redundancy.

10.3.4.5 Orbit Usage

Clusters are equally applicable for operations in geostationary, low earth or highly elliptical orbits and provided due account are taken of the specific orbital environments the hardware can be identical. This ability would provide a low-cost solution to operation in both GEO and TUNDRA if this was required.

A physically linked set of satellites, as in some forms of cluster, may rate as only one satellite for orbit location purposes. A particular case of interest is a small cluster of three satellites linked by a tether, with one satellite at geostationary altitude while the other two are located above and below along an extended earth radius line. This arrangement permits three satellites to operate from essentially one geostationary location. This would be of particular value at the congested locations.

Cluster systems may also consist of spacer.^ft having operational roles in different disciplines. These could include communications, surveillance, meteorology and elint functions in either a mixture of disciplines or as separate clusters and have a cost advantage due to commonality of hardware.

10.3.4.6 Costs

The main cost drivers of a satellite system are as follows:

. Development cost of platforms and payloads and
. The recurring costs.
. Launch costs.
Operating costs.

a) Development costs.

These costs arise whenever a new system or changes to an existing system are required. For SATCOM systems which already contain a high infrastructure, investment changes can only be economically implemented gradually over a long period. A case in point would be a transition from a largely SHF system towards one which may be dominantly EHF. A cluster based system would allow replacement of SHF with EHF in small increments in harmony with slow changes to the ground segment. The present system of procurement and changeover is illustrated in Fig. 10.9 which requires three new spacecraft and payload designs and developments in say a 21 year period. The detterent designs differ in the ratio of capacity provided in each frequency band and is illustrated by the size of the box in the figure

Figure 10.10 Shows the s ı ne conversion from SHF to EHF using a cluster with shows only one development at the outset. Assuming the development costs for each design are similar then over the 21 year period, development costs for the cluster implementation are one third of the current conventional method.

b) Recurrent Costs.

The ratio of recurrent to development cost for satellites lie in the range 1.4 and 1.2.5 dependant on complexity and the extent of new techn logy in the design. The effect on recurrent cost of amortising the development cost over the number of spacecraft produced, becomes smaller as the number of spacecraft rises. Equally the effect of amortising the cost of tooling and test equipment used in manufacture on recurring costs decreases as the number of identical modules used in the bus and payload rises. The design philosophy for closters systems automatically tends to reduce the effect at spacecraft level but the effect would not be insignificant unless the number of spacecraft was greater than about 16. Few if any SATCOM systems would use so many identical spacecraft ever for the 21 years used here for comparison purposes. However numbers of the identical subsystems used in the satellites of a cluster system could easily exceed this with only 4 spacecraft of the same design. The modular approach to design at subsystem level reduces amortisation costs. Fig. 10.11 shows effect of amortisation for development-to-recurring cost ratio against numbers of identical objects. The figure shows the number of identical satellites which are required to amortise with the percentage increase given by each curve.

For the purpose of evaluation of different SATCOM system implimentations a ratio of 3 1 will be used.

c) Launch Costs

The system cost due to launches is a function of the cost of the launch vehicle, the launch operations and the number of launches in the assumed system life time. The greater the choice of launcher which can be utilised, the lower the costs will be at any given time. This is due to the high degree of competition which exists and to the wide range of choice of dedicated and piggy back launch slots which would be available to match the requirement. A cluster system may be implemented with only a single replacement satellite payload when needed. This flexibility gives a good match to the required need as it arises and leads to lowest cost. The number of launches required for a cluster system appears inherently greater than for a conventional approach, but over a period of say 21 years the total mass pl ed in orbit will be smaller due to the high level of redundancy available and the lower masses required for satellite replacement.

d) Operations Costs.

These costs depend on the degree of autonomy within the satellite system. If autonomy is low, as in current systems then these costs will be proportional to the number of satellites in orbit which require monitoring and control of routine functions. A ciuster system would then be more expensive to operate. It is however expected that in the period 2000 to 2300 a high degree of autonomy can be implemented and a cluster can be considered simply as a large satellite. The operations costs for repair should not be greater than for the same degree of repair of a non distributed system.

10.3.5 Degree of Implementation.

The foregoing is intended to identify the potential of cluster Systems in their application to SATCOM and other space systems of military interest. The extent to which all or some of the attributes can be obtained is clearly a function of available finance as well as the weight put on operationally desireable functions. At almost any level more capability on a through life costing basis can be obtained by clusters systems than by more convential approaches. Some comparisons are made to illustrate this premise in Section 12.2 -Evaluation of Candidate Architectures.

10.4 TETHERED SATELLITE SYSTEMS

Various concepts of the application of tether in space have been put forward in the period 1960 to 1980. These ranged from systems which conserved or generated propulsive power for satellites, provided ULF/ELF Antennas in space, to a means of locating more than one satellite of a single geostationary arc position. A brief history of development of tether in space is given in Ref. 10.5. Todate no practical implementation of tether systems have been produced or used largely due to material mass/strength considerations.

Of the systems proposed, that having the most promise and application for satellites communications is a geostationary system of tethered satellites aligned along an earth radius line. The satellites are located above and below the geostationary in a balanced way and stabilized at the altitude of the geosynchronous orbital speed. It is claimed that the number of satellites which could be located at a single orbital allocation point could be increased by a factor of thirteen and in time could alleviate geostationary congestion at popular locations (see Reference 10.6).

An extension of the system is to use the tether link between satellites to either centralize certain resources or to share such resources as power and redundancy among satellites to provide greater service availability for the set (see Reference 10.7).

Further advantages could be derived from a two satellite tethered system in which the tether length would be several thousand kilometers in length. The lower of the two satellite systems would then have lower propagation delay and need less prime power, and hence be cheaper.

The upper satellite could sir / be a counterbalance or have a role such as data relay where propagation delay is of small consequence

The prospect of implementing such systems is some time hence and requires resolution of problems of tether material mass/strength and immunity to micro-meteor damage, deployment and stability during both deployment and operational modes.

The foregoing applies to which are flexible, of significant mass and where the link between the satellites imposes or makes use of stress between satellites. Other times where this is not so, for example optical fibre links may however be implemented to carry data or control signals between independently controlled and stabilized satellites (see section 10.3 on Cluster Systems).

10.5 LIGHTSAT AND PROLIFERATED LEO SYSTEMS

This is not as much a national program as it is a philosophy or a movement. This philosophy emerged several years ago and has undergone several evolutions since then. It has been known under various names such as CheapSat, SmallSat, and Multiple Small Satellite Program (MSSP) [10.8]. Over the years DARPA has been the primary motivating force for the LightSat approach, although the philosophy is now rapidly gaining adherents in the military services and the commercial sector as well. The United States Congress recently increased DARPA's funding request for their LightSat Program.

This philosophy originally sprang from the work on packet switched data network at DARPA (ARPANET). One early concept was a worldwide packet switched network consisting of several hundred (220 was one number considered) satellites interconnected with each other by crosslinks and with ground nodes by uplinks and downlinks. The satellites would be very low in cost, hence in low orbit. The system would be robust because of the large number of alternate paths that would exist between any two fixed ground nodes. It would be survivable because of the proliferation of satellites and finally it was hoped that service rivaling that of existing satellite systems could be provided at lower cost.

The main technical issues were: (1) The complexity of algorithms needed for the dynamic-node network of satellites in different orbits. (2) the difficulty in predicting exactly how well such a system would perform, and (3) the cost could not be made competitive with higher orbit systems, which required many fewer, albeit more expensive, spacecraft to obtain equivalent coverage.

Missions other than conventional space communications have also been considered as applications for the LightSat philosophy. Of these, store and forward communications and signals intercept appear particularly attractive. These are missions which can be carried out with simple payloads and for which the limited coverage of low orbits is of little concern. Other applications might deal with the reconstitution of essential communications and reconnaissance capabilities after a nuclear strike.

The LightSat philosophy today encompasses several related effort. First there is the DARPA program of the same name. This includes efforts to develop low cost launch vehicles such as the Pegasus, which is being designed by Orbital Sciences Corporation to be launched from an aircraft into low earth orbit. The Navy is pursuing LightSat applications under the name SPINSAT and will test a Pegasus launched payload this year (See Section 8.13).

The Air Force has a similar effort going under the name "Reserves". The Army is looking at a number of applications as well. The LightSat concept was recently embraced as a part of the future military communications satellite architecture for the United States.

It is foreseen that the LightSat approach will make a significant impact on the future utilization of space in both commercial and military applications. It is unlikely, however, that LightSat will completely supplant current satellite systems or their lineal descendants. LightSats are seen as an augmentation and expansion of our present uses of space but not as a total revolution [10.9].

10.6 SUMMARY EVALUATION

10.6.1 Proliferated (Multiple) Satellite System (MSS) in LEO

This is basically a global system and as such would have frequency coordination problems (interference) and high cost. Transition problems would almost certainly occur. The advantages, however, of the concept are that there is a high degree of connectivity which would give good survivability with graceful degradation. However, several of the features required need to be analysed and tested.

10.6.2 Tethered Satellites

The geostationary tethered satellite system expands the geostationary orbit resource from a one dimensional arc to a two dimensional disc. The tethered satellites, each several thousand kilometres apart and aligned along a local vertical, are stabilised at the allitude of the geosynchronous orbital speed. When this system is applied to communications systems, it is found that the number of satellites and the communications capacity can be increased many fold compared with a conventional geostationary orbit satellite system. This concept and variations thereof are being studied by a number of organisations and the problems of stability and tether materials still require much analysis and development. Despite the promise this concept is not considered to solve any of the NATO problems effectvely and economically. e.g. polar coverage and physical survivability. and is not considered further.

10.6.3 Inclined Elliptical Orbit Types

Within the general class of inclined elliptical orbits those having 63.4° inclination have been identified as the only ones suitable for NATO use (see Sec. 10.2.3). This inclination prevents the axes of the ellipse from rotating in the plane of the orbit and is thus inherently stable. The following sub-classes are of potential use for the coverage of the NATO area including in particular the polar region.

a) LOOPUS.

This class of orbit provides continuity of service with a minimal tracking. Satellites are seen at high elevation angles which ensure a minimum of environmental effects (atmospheric attenuation and ease of frequency co-ordination). The system is designed to give full coverage to the northern hemisphere using between six and nine satellites. For NATO, however a variant of the system using three satellites would suffice.

b) MOLNIYA.

Two variants of this class of orbit are considered:- 12 hour and 24 hour orbits. To provide continuous coverage for NATO three satellites in 12 hour and two satellites in the 24 hour case are required. Despite the advantage in terms of the number of satellites the 24 hour system implies higher launch costs. longer propagation delays and increased path loss which greater transmit power. Also unlike the 12 hour system it does not offer the possibility of continuous service with a single RF head ground terminal. It should be noted that the 12 hour case can be a variant of LOOPUS.

c) TUNDRA

This class uses two 24 hour orbits with an ellipticity of around 0.4 to provide full coverage of the NATO area. It suffers from the same disadvantages as the 24 hour MOLNIYA case in (b) above but has further disadvantages in an even higher launch cost and

greater path loss. It is however less physically vulnerable to attack. Two RF heads are required in ground terminals to avoid a gap in service at satellite changeover.

10.6.4 Geostationary.

(as NATO IV) This type can be multi-frequency and for NATO coverage (without the polar region) only one operational satellite is needed. Its main disadvantages are the high launch cost, limited coverage and low elevation angles of access. However as far as NATO is concerned, a position in the geostationary arc is a valuable resource and should not be given up lightly.

10.6.5 Systems of Satellites in Combinations of Orbit Type.

For this class of system, satellites are used in geostationary orbit and supplemented by satellites in other orbits to provide polar coverage.

a) Geostationary plus Circular Polar.

This system uses a spacecraft with a dual frequency payload in geostationary orbit and six single frequency spacecraft in a circular polar orbit at an altitude of 5000 km. or higher. The polar satellites could have either SHF or EHF payloads. Coverage is not fully continuous since as soon as the operational satellite disappears over the horizon the ground station must slew about 60° to acquire the next satellite as it comes into view. This could however be acceptable for the traffic conditions expected ta apply in the polar regions. The polar satellites would form an independent system of satellites and would have the merit of matching the polar satellite design to the special user needs of the region. This could lead to relatively simple satellites which could make them affordable despite the large number of satellites required. The polar satellites could be linked to the geostationary system using either inter-satellites links or terrestrial inter-connection.

b) Geostationary plus 24 hour MOLNIYA.

Systems of this class use satellites in both GEO and MOLNIYA orbits. The system may be implemented by:-

i) A dual frequency spacecraft in GEO with two single frequency (EHF or SHF) spacecraft in MOLNIYA orbit.

ii) One single frequency spacecraft EHF and one single frequency SHF spacecraft in GEO together with two single frequency EHF or SHF spacecraft in MOLNIYA orbits

The first of these provides the lower cost of the two options but requires two different spacecraft designs to implement it and two different designs to replace failed or spent satellites. The second option has a higher initial cost to implement but the same design can be used for both orbits provided the implications of the two orbit types are taken into account in the design phase. The spares needed for failures in either orbit would be the same. A further advantage of the second system (b) is the ability o mix the ration of EHF and SHF satellites when traffic requirements change during the system lifetime. Although the 24 hour MOLNIYA orbit is proposed and not the 12 hour, the lack of total continuity of cover at handover between spacecraft is considered acceptable since the spacecraft in inclined orbit primarily serve the polar region. It is worth noting that the spacecraft in MOLNIYA orbits provide supplementary capacity at EHF or SHF to that of the geostationary spacecraft in the current NATO region.

10.6.6 CLOUDSAT.

This system makes use of several satellites which work in the receive mode and are spatially separated so that one jammer cannot jam more than one satellite at a time. These receiving satellites transmit via cross links to a common downlink. The system uses as many satellites as are necessary to gain the required jamming advantage. The use of many satellites also provides a measure of in-orbit redundancy which is additional to the more normal internal redundancy which one might have. The Cloudsat concept can be used in any orbit and for NATO coverage there could be, for instance, two clusters in two or three MOLNIYA orbits. Costs would be dependant on the number on the number of satellites in the system.

10.6.7 MEWS.

This system makes use of satellites in a cluster formation. The satellites in the cluster can be single or multi-frequency with combined or separated receivers and transmitters. The satellites in the cluster are inter-connected by means of laser or radio links to enable centralisation or redistribution of such services as attitude control, adaptation to traffic needes and the provision increased redundancy. The system can also provide a means for in-orbit refuelling and economical repair and evalutionary growth. With sufficient spatial separation among elements of the satellite cluster similar ECCM gains as those offered by Cloudsat can be obtained.

10.6.8 Number of Satellites.

The number of satellites required for full coverage and service continuity is given in Table 10.1. The table also shows the number of satellites needed for a 21 year system life considering:-

(a)the features that are required in a future NATO system.

(b)the characteristics of the various candidate architectures.

N.B. The data in Table 10.1 is based on the following assumptions:-

1) Availability of service required of the systems = 0.95

2) Satellite reliability (7 years):-
Dual frequency= 0.61 Single frequency (EHF) = 0.72. (SHF) = 0.75

3) The 21 year period is made up of three independent 7 year periods

4) Satellite design is independent of orbit type for cost effectiveness.

5) For comparison all satellites are assumed to have 7 year lives

6) Satellites in any option can operate in one or two frequency bands (SHF and EHF) See section 11.2

7) For other assumptions see section 11.2

10.7 CANDIDATE ARCHITECTURES

The group has concluded that the following architectures all have their own relative advantages and disadvantages. All appear to need similar levels of R&D, and are considered technically feasible within the time frame 2000-2030. The actual choice of an architecture to implement, will need to take account of actual requirements, which can only be known with certainty nearer the time of implementation. All of the systems can be designed to have similar communications capacity, levels of availability, and survivability to electronic and physical threat. For the purposes of comparison each system will be assumed to have the same communications capacity and AJ capability, have EMP protection and be hardened to the same degree against collateral nuclear effect. The satellite communications features would be consistent with the desired characteristics outlined in

Section 10.1 and as defined in Section 11.1.

Option 1.

Operational satellites in geostationary orbit supported by inorbit spares to ensure adequate continuity of service. Area coverage will be as for NATO IV. For polar coverage and communications to submarines NATO would lease capacity from USA/DND. In cases of premature or catastrophic failure NATO would gain access to other national SATCOM systems via interoperability agreements. A variant of this architecture could utilise Cloudsat or MEWS features to some degree. This system is regarded as baseline for comparison with other architectures.

Option 2.

Operational satellites in 12 hour Molniya orbits supported with in-orbit spares to provide service continuity. This architecture would provide full NATO area (including Polar) coverage except for submarine communications. Capacity for this aspect would be leased from the USA/DND. Inter-operability agreements would apply as for option 1. Cloudsat and MEWS features could also apply to this architecture but linkages between the three orbits would not be effective.

Option 3.

Operational satellites in both geostationary and 24 hour Molniya orbits with in-orbit spares for service continuity. Those in geostationary orbit would provide coverage as for NATO IV while satellites in the two 24 hour Molniya would provide cover of the polar regions. Submarine communications would be obtained from the USA as in options 1 and 2. Cloudsat and MEWS features could also apply in this option.

NOTE 1. Laser communications to submarines may become available in the plus 2000 time frame but would depend mainly on success in US R &D. Current opinion is that a technological breakthrough is required and even if this occurred a separate and independent Lasercom satellite would be necessary due to the large equipment mass and power generation system required. Should this be the case NATO is unlikely to afford such a system of its own.

NOTE 2. All three options may be implemented with spacecraft with either single frequency payloads or multi-frequency payloads.

NOTE 3. Comparative evaluations made in section 11.2 include versions of the options which include inter-satellite links used for both reliability enhancement and limited Cloudsat type spatial AJ purposes.

10.8 REFERENCES

[10.1] "Rocked Propulsion and Spacecraft Dynamics" Section 18.6 (Book-author and publisher unknown)

[10.2] Masterman, P.H. "Initial Considerations for a Post-2000 NATO SATCOM System", SHAPE Technical Centre TM-886, 1990 (NATO Secret)

[10.3] Collins, J. et al., "Tactical Military Communications by Satellite Relay at High Latitudes", AGARD Conf. Proc. No. 344 on Space System APplications to Tactical Operations, 1983, pp 10-1 to 10-15

[10.4] Ahmed, M.J., "Tactical EHF Satellite Communications for NATO Supplying Non-Geostationary Orbits", AGARD Conf. Proc. No. on "Tactical Applications of Space Systems", Colorado Springs, 16-20 Oct. 1989.

[10.5] "Historical Education of Tethers in Space", I. Bekey, NASA Headquarters Washington DC - NASA/AIAA/PSN International/Conference on Tethers in Space 17-19 Sept. 1986, Arlington VA.

[10.6] "Geostationary Tether Satellite System and its Application to Communication Systems", Tetsoo Yasaka and Takeshi Hatsuda-IEE Transactions on Aerospace and Electronic Systems Vol. 24 No. 1 January 1988.

[10.7] Tether in Space Handbook - NASA Office of Space Flight: Advanced Programs; CODE MT August 1986

[10.8] "Low Altitude Multiple Small Satellite System", Proc. IEEE, Jan. 1987

[10.9] D.L. Cromer, "Buck-Rogers - What can you do for us?" AGARD Highlights 90/1

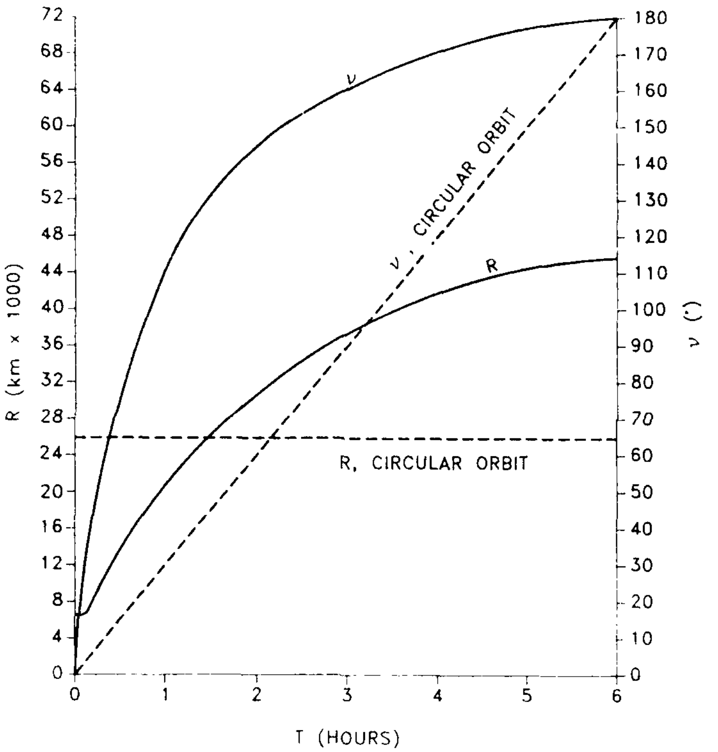


Fig 10.1 Time-dependence of radial vector for a satellite in a highly-elliptical 12-hour orbit (Perigee altitude 100 km)

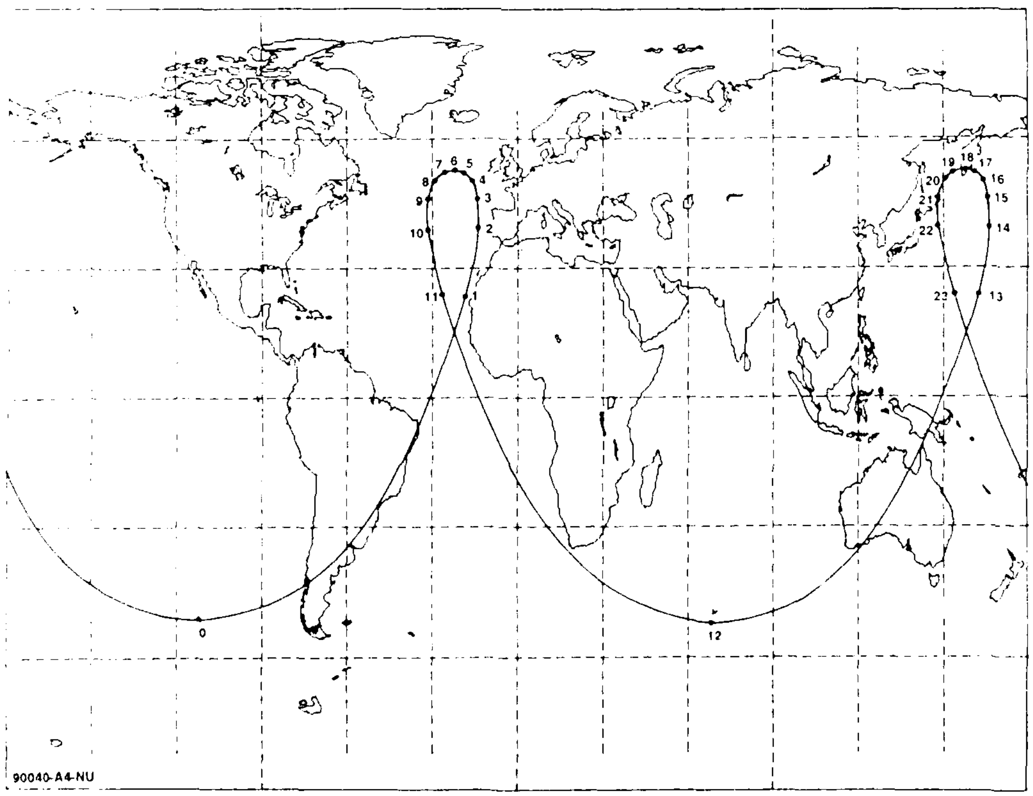


Fig.10.2 Ground trace of satellite in 12-hour orbit with Perigee altitude 1000 km and inclination 52°

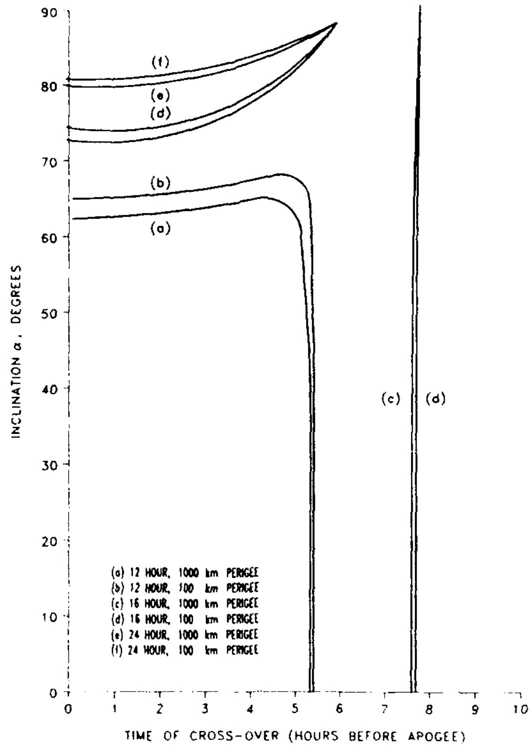


Fig.10.3 Timing of ground-trace cross-overs for highly elliptical orbits

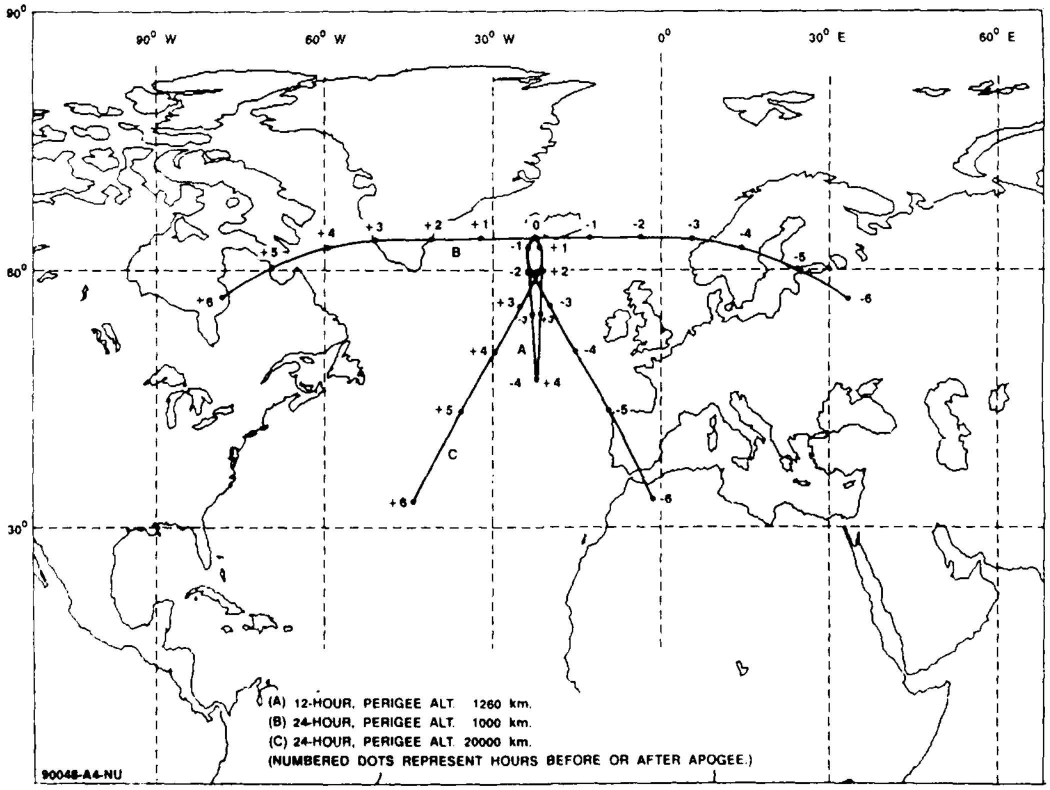


Fig.10.4 Ground traces of satellites in elliptical orbits with 63.43° inclinations

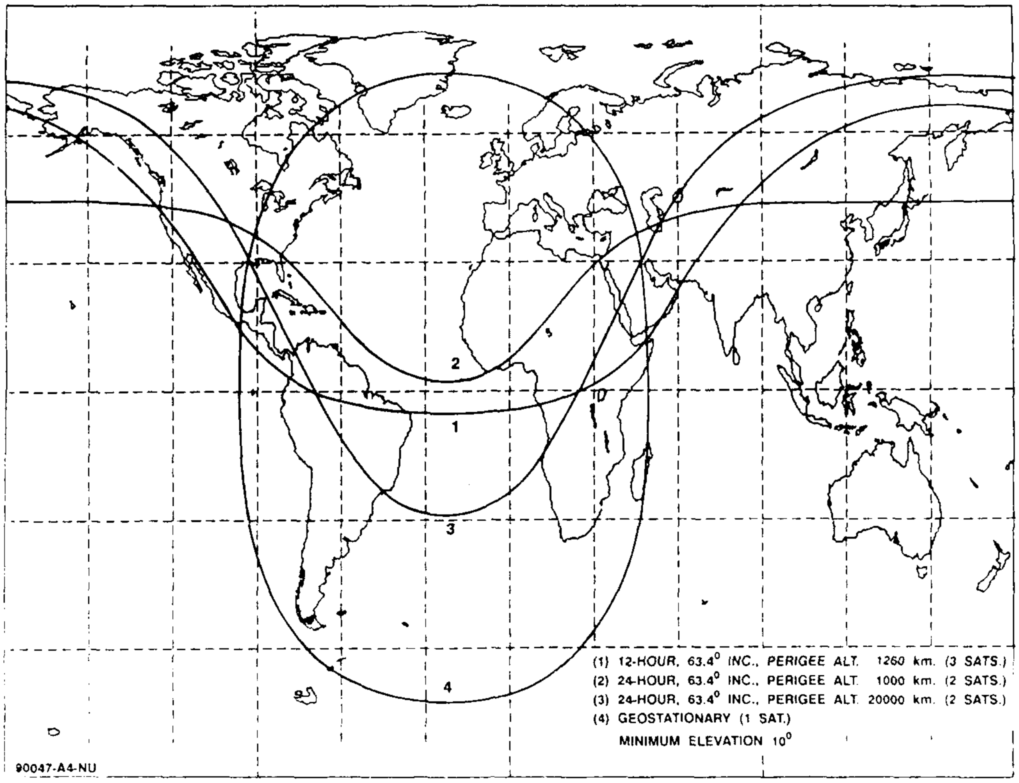


Fig.10.5 Coverage limits for various satellite systems

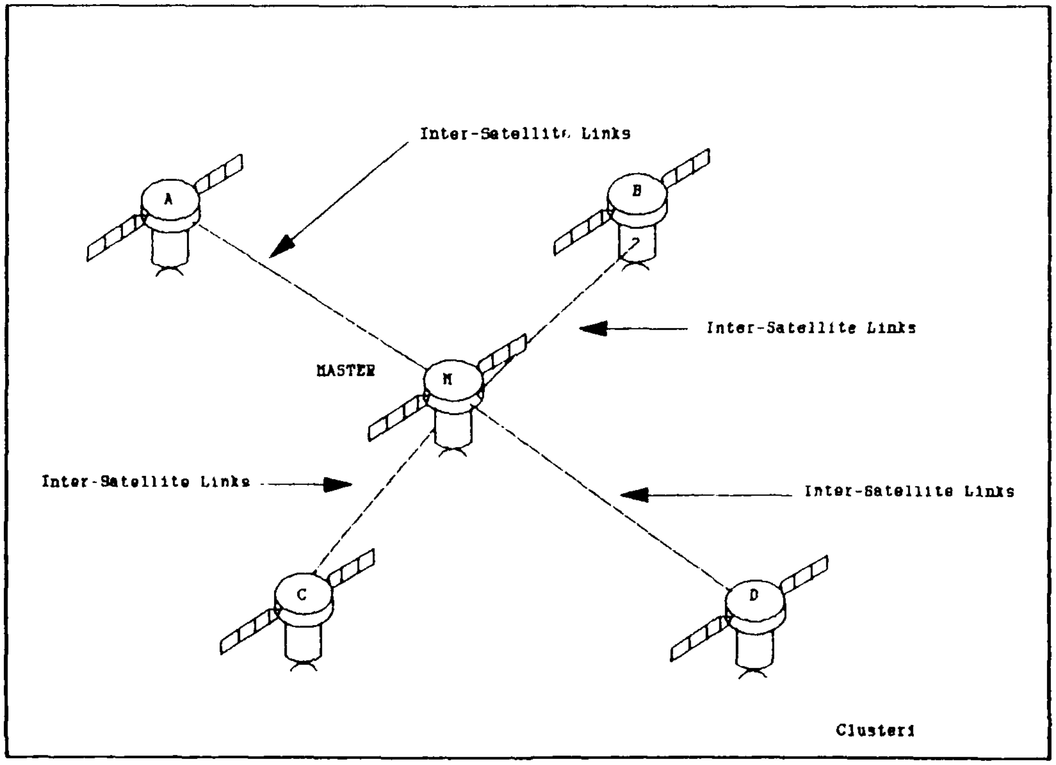


Fig.10.6 Cluster system

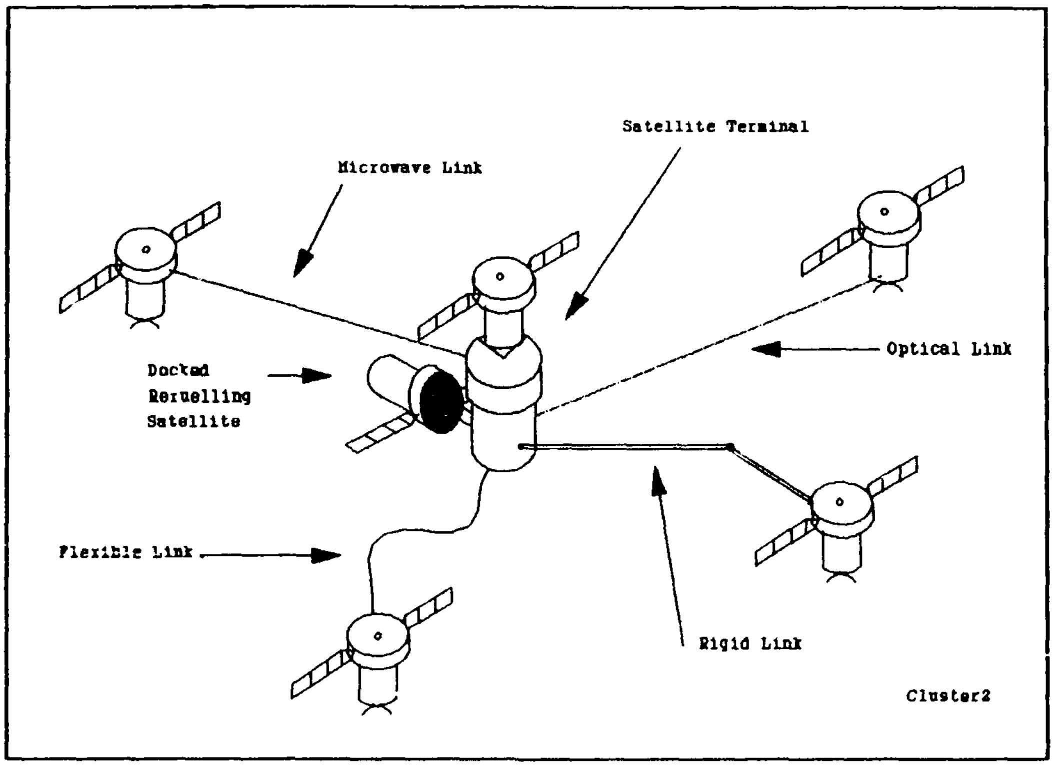


Fig.10.7 MEWS system

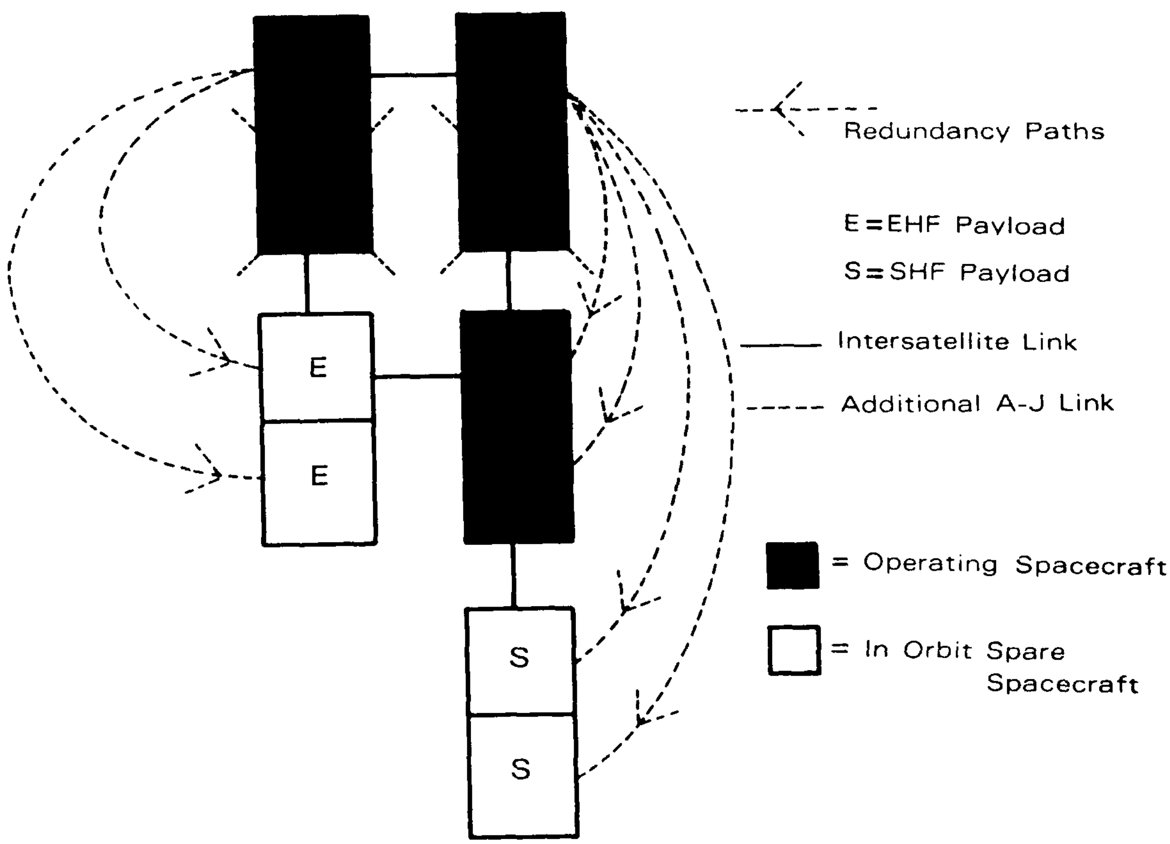


Fig 10 8 Cluster configurations redundancy paths using a mixed frequency payload

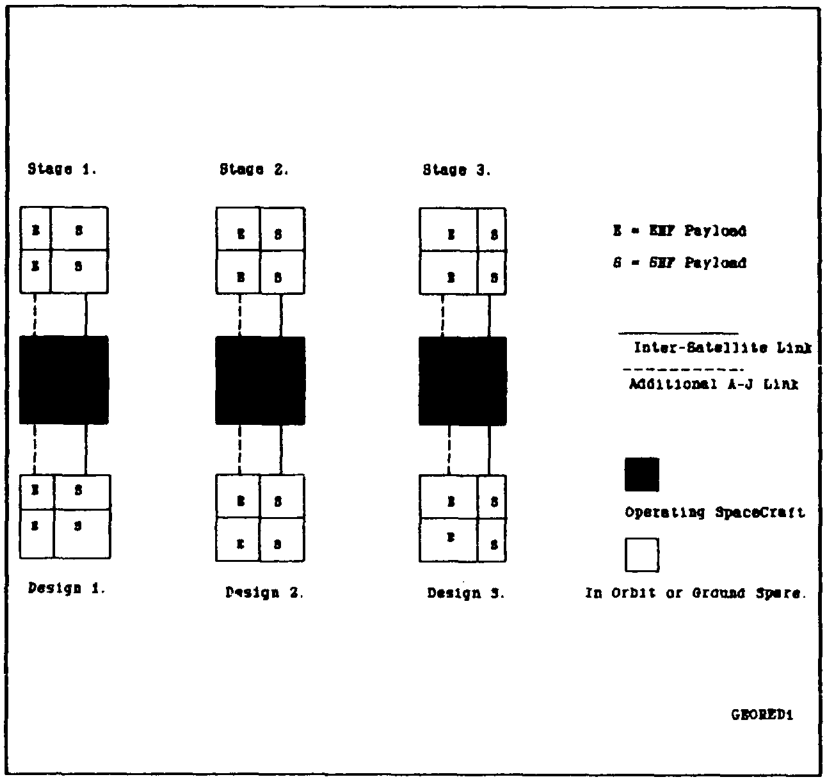


Fig.10.9 Deployment of dual frequency redundant spacecraft in geo-stationary orbit

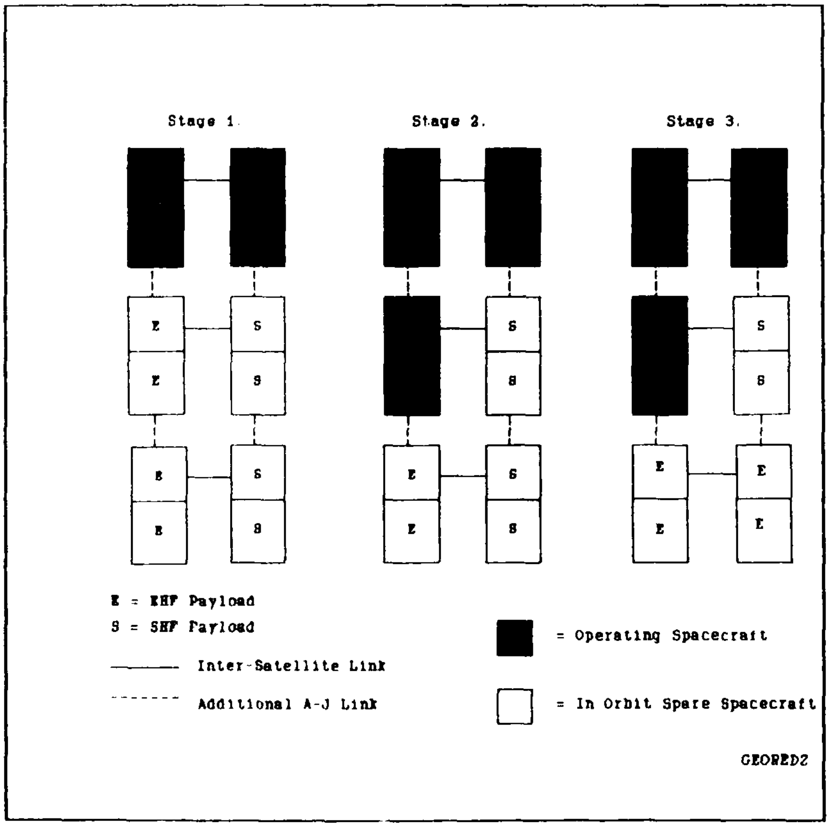


Fig.10.10 Deployment of single frequency redundant spacecraft in geo-stationary orbit

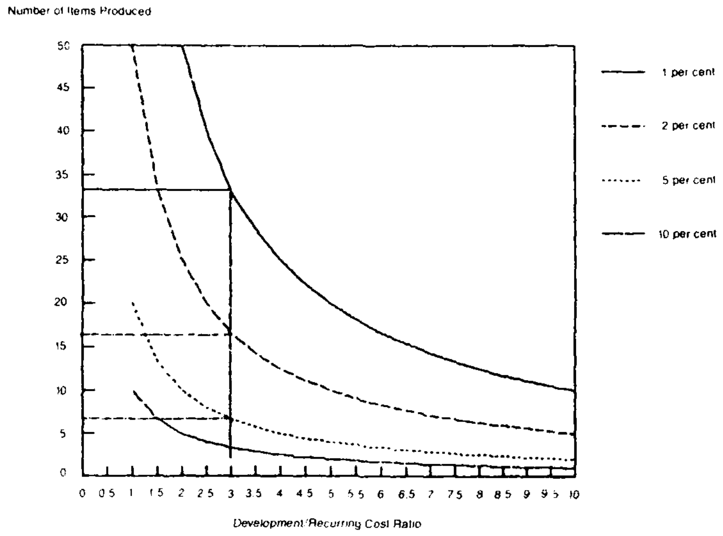


Fig.10.11 Development to recurring cost versus number of items produced

Table 10.1
Number of satellites required for different SATCOM architectures

Features Architecture	No of Active S/C for full continuous Coverage	Comment	No of S/C for 7 years	Total No of S/C for 21 years
(a) Proliferated LEO	240		>240	>500
(c) Inclined Elliptical Orbit (c1) LOOPUS (c2) MOLNIYA (c3) TUNDRA	 9 3 x 24 - hrs 2 x 24 - hrs 2 x 24 - hrs		 27 9 6 6	 81 27 18 18
(d) GEO	1	Baseline	3	9
(e) Systems of Satellites in more than one Orbit (e1) GEO + POLAR	 1 GEO + 6 POLAR	 S/C in GEO are dual frequency, single in Polar Orbit	 3 GEO + 18 Polar	 9 GEO + 54 Polar
(e2) GEO+24-hr MOLNIYA (a) Dual freq S/C in GEO+ single in Inclined Orbit (b) Single freq in both orbits	 1 GEO + 2 inclined 2 GEO + 2 inclined	 GEO S/C have EHF and SHF Inclined EHF or SHF GEO S/C are mix of EHF and SHF Inclined EHF or SHF	 3 GEO + 6 inclined 6 GEO + 6 inclined	 9 GEO + 18 inclined 18 GEO + 18 inclined
(f) CLOUDSAT vith 10 dB ECCM advantage relative to (e2)b	Receive S/C 20 GEO, 20 Inc. TX S/C: - 2 GEO + 2 Inc	Numbers will depend on jamming threat at the time	60 GEO 60 inclined 6 GEO 6 inclined	180 GEO 180 inclined 18 GEO 18 inclined
(g) MEWS	The basic concept applies to all cases c thro' (f)			

APPENDIX 10A

CANDIDATE ORBITS FOR A POST - 2000 NATO SATCOM SYSTEM

1. BACKGROUND

The purpose of this Appendix is to identify possible systems in terms of numbers of satellites and types of orbits which would provide continuous coverage of at least the region currently of interest to NATO i.e. that enclosing the locations of the 21 large static SGTs. Eventually these candidate systems will be evaluated for a post-2000 SATCOM application, but this cannot be done until the on-going threat assessment is complete and future operational requirements are more clearly established.

The candidate systems fall into two distinct classes:

(i) Those in which a single satellite is visible from all SGTs at any one time. Except in the special case of a geosynchronous (24 hour) orbit enough active satellites must be provided so that, as seen from any SGT, each satellite rises well before the previous one sets, giving time for all circuits to be transferred from one satellite to another. This type of system supports real-time circuits as in the current NATO SATCOM system.

(ii) Those in which every SGT can see at least one satellite at any one time, but where in general all SGTs do not see the same satellite simultaneously. This type of system is suitable for store-and-forward message traffic but if inter-satellite links are used it can also provide near-real-time (e.g. voice) circuits.

2. Orbits for Simultaneous Visibility

In this first category the following possibilities are considered:

(i) Equatorial circular orbits

(ii) Polar circular orbits

(iii) Inclined circular 24 - hour orbits

(iv) Inclined elliptical 12 - hour orbits

(v) Polar elliptical 12 - hour orbits

(vi) Inclined elliptical 16 - hour orbits

(vii) Inclined elliptical 24 - hour orbits

2.1 Equatorial Circular Orbits

For an orbital radius of 42,240 km (altitude 35940 km) a satellite in a circular equatorial orbit has a period of exactly 24 hours and hence appears to maintain a fixed position in the sky. For any other radius the satellite will appear to an observer to drift across the sky and hence several satellites will be needed to provide a continuous service. The number of satellites required (assumed equally spaced around the orbit) can be found by determining the limits of the angular range through which a satellite can move while still being visible from all SGTs simultaneously at an acceptable elevation. Expressing this arc length as a fraction of a full circle, then taking the reciprocal, gives the minimum number of satellites required.

With reference to Fig 1, an observer at point P on the earth such that the angle between P and the sub-satellite point S, as measured at the centre of the Earth, is θ, sees the satellite at an elevation el such that:

$$r_s \cos(\theta + el) = r_e \cos el \qquad (1)$$

where r_s is the radius of the orbit and r_e the radius of the Earth

For a satellite in equatorial orbit where the sub-satellite point has longitude lg_s and latitude zero, and where the observation point P is at longitude lg_p and latitude lt_p, the angle θ is given by

$$\cos\theta = \cos lt_p \cos(lg_p - lg_s) \qquad \ldots(2)$$

For a given orbit radius r_s the value of θ corresponding to a given minimum elevation el can be found from (1). Substituting for θ and for the observer's position in (2) then allows the two values of lg_s corresponding to minimum acceptable elevation to be found.

By determining the allowed range of lg_s for each SGT, the extent of the range common to all SGTs, and hence the minimum number of satellites required, can be found. In the case of the NATO large static SGTs, it is found that provided the satellite is visible at adequate elevation from F3 (Norfolk), F8 (Carp) and F16 (Bjerkvik) then it will also be visible at adequate elevation from the other terminals. Table 1 shows the allowed range of lg_s for each of these three SGTs for a minimum elevation of 10° and for various values of the parameter r_s/r_e. The range of lg_s for visibility from all SGTs simultaneously is also given together with the corresponding minimum number of satellites. The number of satellites as a function of r_s/r_e is also shown in Fig. 2. It can be seen that common visibility is not possible using satellites below geostationary heigh (r_s/r_e=6.7) and that the number of satellites (except for the special case of geostationary orbit) is very large even for much greater altitudes (e.g. 20 for r_s/r_e=10). In any case the propagation delay associated with these altitudes is unlikely to be acceptable for speech circuits.

2.2. Polar Circular Orbits

As Fig 3 shows, the "footprint" of a satellite in equatorial orbit is not well matched to the distribution of NATO SGTs, all of which lie above 36° N latitude. Hence the attraction of a satellite in polar orbit, whose footprint when over the pole can encompass all the NATO SGTs at a much lower altitude. The minimum altitude for visibility above latitude lt_{min} is obtained from (1) with θ equal to $\pi/2 - lt_{min}$. This corresponds to r_s/r_e = 2.25 for 10° minimum elevation.

For greater altitudes, common visibility down to latitude lt_{min} is achieved when the satellite is within an angle ψ of the pole such that $\pi/2 - lt_{min} + \psi = \theta$ in (1). The minimum number of satellites for continuous coverage is then π/ψ. This number is shown in Fig 2 as a function of r_s/r_e. It is evident that while this is a much more practical proposition than the equatorial (non-geostationary) case, the number of active satellites required is still quite large (e.g. 11 for geosynchronous altitude).

2.3. Inclined Circular 24-Hour Orbits

Fig. 4 shows the area of coverage of a geostationary satellite located at 18° W longitude, for a satellite elevation of at least 10°. The extent of coverage for elevation angles down to zero is also shown. It is clear that coverage extends to at most latitude 71° for 10° elevation and 81° for zero elevation. Hence there is interest in the use of inclined geosynchronous orbits to extend coverage to polar regions. Whereas a geostationary satellite apears to remain at a fiwed point in the sky, a satellite in inclined geosynchronous orbit will appear to describe a figure-of-eight path once per day. At high latitudes the satellite will dip down below the horizon for some period each day so a system of two

or more active satellites in orbits whose nodes are equally spaced around the equator will be required to maintain a continuous service. The area in which continuous service can be obtained from a single satellite is reduced as compared to the geostationary case.

For a satellite in a 24-hour circular orbit with inclination α whose ascending node is at longitude lg_{nd}, the latitude and longitude of the sub-satellite point at time t hours after passing through the node are given by:

$$lt_s = \sin^{-1}(\sin \nu \,.\, \sin \alpha) \quad ...(3)$$

$$lg_s = lg_{nd} + \sin^{-1}(\tan lt_s/\tan\alpha) - \pi t/12 \quad ...(4)$$

where $\nu = \pi.t/12$ in this special case.

The elevation of the satellite as seen by a observer at longitude lg_p and latitude lt_p can be shown to be:

$$el = \tan^{-1}((r_s \cos\theta - re) / (r_s \sin\theta)) \quad ...(5)$$

where

$$\cos\theta = \cos lt_s \cos lt_p \cos (lg_p - lg_s) + \sin lt_s \sin lt_p \quad ...(6)$$

It will be noticed that (5) is a re-arrangement of (1) and that (6) is a generalisation of (2).

The above expression allows the elevation of a satellite with a given orbit inclination as seen from a particular point on the earth to be calculated as a function of time. Hence the length of time that the satellite's elevation exceeds a given value can be found. By repeating this for a number of observation points it is possible to derive contours showing the portion of the earth's surface in which the satellite is seen for at least N hours per day above a given elevation. Fig. 5 shows the contours for 8, 12 and 24-hour coverage at 10° elevation or more for a satellite whose orbital inclination is 20°. The 8-hour coverage extends to about 80° latitude but the 24-hour coverage now encloses only 6 of the 21 NATO SGTs. Hence the advantage of higher latitude coverage is gained at the expense of a system with three active satellites instead of one, and with the need to transfer most inter-static circuits from one satellite to another at least twice a day. Such handover would imply temporary outages unless the affected SGTs were double-headed

2.4. Inclined Elliptical 12-Hour Orbits

These orbits have been used by Soviet "Molniya" satellites for many years. At apogee the satellite is approximately at geosynchronous altitude whereas at perigee its altitude may be only a few hundred kilometers. Near apogee the satellite movement more or less keeps pace with the rotation of the earth for several hours and thus it maintains a fairly constant "footprint" during this period. Continuous coverage can therefore be provided using only a small number of active satellites. Each satellite provides coverage of the service area at alternate apogees, i.e. every 24 hours. In principle the inclination of the orbit is not very critical, but in practice an inclination of 63.4° is often used since such an orbit is found not to precess, and hence less fuel is required to maintain the orbit parameters. A major advantage of the Molniya orbit is that the apogee boost motor used to circularise the orbit of a geosynchronous satellite is not required: typically this accounts for about half the mass of the satellite at launch

It can be shown from Kepler's third law that the orbital period T of a satellite is given by:

$$T = 2\pi\mu^{-1/2}a^{3/2} \quad (7)$$

where a is the semi-major axis of the orbital ellipse and μ = 3.986094×10^5 km^3/sec^2 for the earth. From this it follows that a = 42,240 km for a 24 hour orbit and 26,610 km for a 12hour orbit. Derivation of the satellite's position relative to perigee in polar coordinates r_s, ν as a function of time is slightly awkward but can be done by introducing a variable x such that $dx/dt = \sqrt{\mu}/r_s$. It can be shown from consideration of Kepler's laws that:

$$r_s = a(1 + e \sin ((x + c_o)/\sqrt{a})) \quad ..(8)$$

Here e is the eccentricity of the orbit and c_o is an arbitrary constant. By substituting $\sqrt{\mu}$ dt/dx for r_s in (8) and integrating w.r.t. x we obtain:

$$\sqrt{\mu} = ax + ae\sqrt{a} \,(\cos ((x + c_o)/\sqrt{a}) - \cos c_o/\sqrt{a}) \quad ..(9)$$

For a given value of ν, r_s can be obtained from the equation of the ellipse, viz.:

$$r_s = p/(1 + e \cos\nu) \quad ..(10)$$

$$\text{where } p = a(1 - e^2)$$

From this the value of $x + c_o$ can be obtained using (8). Substitution of this value in (9) gives the corresponding value of t plus a constant term. By repeating the calculation for another value of ν the constant term can be eliminated and the time taken for the satellite to move between the two values of ν can be found. Fig. 6 shows how ν and r_s vary with t in the case of a 12-hour orbit whose perigee altitude is 1000 km, corresponding to an ellipticity e = 0.726. Values for a circular orbit are shown for comparison. The relatively slow change in the position vector near apogee can be clearly seen.

Equations (3) and (4) are quite general and allow the sub-satellite point to be determined as a function of ν regardless of orbit. They have been used to produce the ground trace shown in Fig. 7, for a satellite in a 12-hour orbit with perigee at 1000 km altitude inclined at 52°. The first apogee is at 22° W. Points on the ground trace indicate the time in hours since the initial perigee. It can be seen that for at least four hours either side of the first apogee the satellite is well placed for coverage of the NATO area. A system with three active satellites in such orbits could provide continuous service to NATO.

In the above example each circuit would suffer an outage every eight hours unless each SGT had a second RF head which could acquire the rising satellite and set up a link to each of the other SGTs via that satellite before transferring traffic from the setting satellite. Whereas tactical terminals could probably tolerate such outages, (there would be considerable flexibility in the timing of transfers), they would probably be unacceptable for inter-static links. As the cost of double-heading the large SGTs would be very high, it is worth considering a way in which the ground trace crosses over itself some 15 minutes before and after perigee. It is found that orbits of this type have 0, 1 or 2 cross-over points depending on the inclination of the orbit, and that the timing of the cross-overs depends strongly on the inclination. Thus by careful choice and maintenance of orbital inclination it should be possible to make a cross-over occur exactly 4 hours either side of apogee, so that the rising and setting satellites appear to pass simultaneously through the same point in the sky at that time and allow all SGTs to transfer their circuits while both satellites are within the antenna beam. At the same time the SGTs cease tracking the setting satellite and begin tracking the rising satellite.

The inclination α required for cross-over at a specified time can be found from consideration of equations (3) and (4). By requiring the longitude lg_s of the sub-satellite point to be the same τ hours before (or after) apogee as it is at apogee it can be shown that:

$$\sin\alpha = \frac{(\sin^2\nu(\tau) - \sin^2\xi(\tau))^{1/2}}{\sin\nu(\tau)\cos\xi(\tau)} \qquad \ldots(11)$$

where $\nu(\tau)$ is obtained as described earlier in this subsection and

$$\xi(\tau) = \pi(1-\tau/6)/2$$

The required inclination as given by equation (11) is shown in Fig. 8 as a function of τ for the case of a 12-hour orbit with perigee altitude of 1000 km. It can be seen that for inclinations less that about 62° the cross-over occurs 30-40 minutes after perigee. For inclinations between 62° and 65° there are two cross-overs, and for inclinations greater that 65° no cross-over occurs at all. For a cross-over 4 hours either side of apogee an inclination of 64.73° is required. This falls in the double cross-over region, where the timing of the second cross-over is extremely sensitive to small changes in inclination. It is interesting to note that this inclination is close to the 63.4° value adopted by Soviet "Molniya" satellites.

The ground trace of the 12-hour orbit with 64.73° inclination is shown in Fig. 9. The two cross-over points are hard to distinguish as the ground trace lies practically along a line of longitude in this region. The boundary of the coverage region is also shown, representing that area in which a satellite is visible at 10° minimum elevation at any time within four hours of apogee. (It should be noted that in Fig. 9 only the coverage area associated with the first apogee is shown, i.e. that at 22° W. There is of course a second coverage area centred on longitude 158° E but this is not likely to be useful to NATO). It is clear that with suitable choice of apogee longitude, the primary coverage area can include the whole territory of all the NATO nations and the entire North polar region. (A corollary of this is that it also inevitably includes a large part of the Warsaw Pact area, and thus offers no *inherent* ECCM protection.)

The inclination required for a conjunction four hours from apogee is also dependent on perigee height. Fig. 8 includes a curve of inclination versus τ for a 100 km perigee altitude as well as the 1000 km case just discussed. It can be seen that the required inclination increases by about 3°. It is evident that there is considerable flexibility in the choice of perigee altitude and thus, for example, the deleterious effects of passage through the Van Allen Belts may be minimised. In the case of a 12-hour orbit it would also be possible to choose the perigee altitude so that a conjunction 4 hours from apogee corresponded to a 63.4° inclination, thus avoiding precession.

The transfer of circuits from one satellite to another must take place during the time both satellites are within the SGT's main beam. To determine the maximum time available for transfer, the angular separation of the satellites when close to conjunction must be calculated as a function of time and compared with the antenna's beamwidth. If the two satellites are seen at azimuth angles Az_1, and Az_2 and elevation angles El_1 and El_2 then the angular separation is given by:

$$\cos\theta = El_1 \cos El_2 \cos(Az_1 - Az_2) + \sin El_1 \sin El_2 \qquad \ldots(12)$$

The elevation of each satellite can be determined from (5) whereas the azimuth is found from

$$\tan Az = \frac{\cos lt_s \sin(lg_s - lg_p)}{(\sin lt_p \cos lt_s \cos(lg_s - lg_p) - \cos lt_p \sin lt_s)} \qquad \ldots(13)$$

For the 12-hour orbit of Fig. 9, (12) has been evaluated in the vicinity of cross-over. The observation point was taken as F1 (Kester) but the result is not expected to depend strongly on this. It is found that the angular separation changes at a rate of 0.54° per minute. Thus for a 12 m SGT with a beamwidth of 0.22° the two satellites will only be within the beam together for 0.22/0.54 minutes, i.e. 24 seconds. In the case of a 2.4 m SGT with a beamwidth of 1.1° this increases to approximately two minutes.

The existence of even this very brief transfer period is dependent on precise maintenance of orbital parameters, in particular the inclination. If the inclination is slightly in error, the ground trace will have a cross-over point somewhat more or less than 4 hours from apogee, but because the satellites are spaced eight hours apart there will no longer be a conjunction at this point. Minimum separation will still occur four hours from apogee but unless that separation is less than half a beamwidth circuits cannot be transferred and the "no-break" nature of the service will be lost. The shift in cross-over time for a given inclination error can be found from Fig. 8. If this shift is t_s minutes then (provided the error is small) the minimum separation can be estimated by determining (from eq. 12) the angle between the position of the rising satellite t_s minutes <u>before</u> cross-over and that of the setting satellite t_s minutes <u>after</u> cross-over. This has been done for the 12-hour orbit of Fig. 9 and its found that the minimum separation is 1.3 times the inclination error. Thus an error of only 0.08° is enough to reduce the transfer time to zero in the case of a 12 m SGT.

2.5. Polar Elliptical 12-Hour Orbits

The system described above using inclined 12-hour orbits clearly provides good coverage of the area of interest to NATO. However, three active satellites are required to provide a continuous service. In this section a variation of the system is considered in which the inclination is increased to 90° and the satellites provide service at every apogee instead of at alternate apogees. In this way it may be possible to provide a continuous service with only two active satellites. Fig. 10 shows the ground trace of a 12-hour polar orbit with 1000 km perigee altitude, centred on longitude 22° West. The elevation of the satellite as seen by an observer at a given latitude and longitude can be *found from (5) where r_s and θ are determined as functions of* time by the method described in section 2.4. This has been done for three stations close to the boundary of the region currently served by NATO SATCOM, viz. Norfolk (F3), Ankara (F6) and Gibraltar (F19). The result is shown in Fig. 11. It can be seen that the satellite is visible at > 10° elevation from all three stations (and by implication from all other large static SGTs) for a period of about 10.4 hours around the first apogee and 5.0 hours around the second apogee. These periods of visibility are separated by gaps of 4.4 hours and 4.2 hours. It can be shown that for continuous service to be possible, the shorter period of visibility must be greater than or equal to the longer gap. This criterion is clearly satisfied. The satellites must be spaced apart by 4.7 hours *in order for the service periods of one satellite to align optimally* with the out-of-service periods of the other. However, the periods when two SGTs can both see both satellites, and hence when circuit transfer is possible, can be fairly short, as is shown in Fig. 12. Here the periods when particular SGTs can see both satellites simultaneously at > 10° elevation are shown in the form of a bar chart. Circuit transfer is possible between two SGTs during periods when their bars overlap. As well as the three SGTs mentioned earlier, F8 (Carp), F15 (Keflavik), F16 (Bjervik) and F21 (Catania) are included as these together with the original three are believed to form a "worst case" set. Particularly critical transfers are F6-F21 and a number of transfers involving F3. However at least 15 minutes is available even in these cases and *the insistence on a minimum of 10° elevation is probably too* conservative.

A two-satellite system of this type therefore appears feasible, but double-heading of all SGTs would be essential, otherwise four outages would occur per day on all links.

2.6 Elliptical 16-Hour Orbits

A satellite in a 16-hour orbit repeats its ground trace every 48 hours and hence has three apogees spaced at intervals of 120°

longitude. The ground trace exhibits cross-overs and hence it is possible to arrange for periodic satellite conjunctions allowing a "no-break" service with single-headed SGTs. Such a system involving six satellites spaced eight hours apart is understood to be under consideration by ESA for a future civil SATCOM system. The timing of the cross-overs as a function of orbital inclination is included in Fig. 8 for cases of 100 km and 1000 km orbital inclination. It can seen that there is always a cross-over soon after perigee, but that a second cross-over only occurs if the inclination exceeds about 70°. This second crossover occurs at most 6 hours before apogee, the 6-hour case corresponding to an inclination of 90°.

Two specific cases have been examined:

(i) Cross-over 4 hours from apogee, inclination 77.45°, perigee altitude 1000 km. (Six satellite system.)

(ii) Cross-over 6 hours from apogee, inclination 90°, perigee altitude 1000 km. (Four satellite system.)

The ground trace for system (i) is shown in Fig. 13. Here the first apogee has been set at 22° West longitude. The limit of coverage associated with this apogee is also shown, and it is clear that this coverage is more than adequate to meet NATO's needs. Similarly the ground trace for system (ii) is shown in Fig. 14 together with the boundary of the primary coverage area. The coverage is more limited than that of system (i) but, again, is quite satisfactory for NATO. By choosing the longitude of the first apogee some 30° further West, the coverage could be made to include all of the continental USA as well as the territories of the other NATO nations.

Possible reasons for choosing a system based on 16-hour orbits in preference to one based on 12-hour orbits include longer "windows" for transferring circuits during conjunctions and reduced sensitivity to inclination errors. These have been investigated using the methods described in section 2.4. For system (i), the angular separation changes at a rate of 0.21° per minute near the cross over point, giving a window of 63 seconds for 12 m SGTs and about 5 minutes for 2.4 m SGTs. The sensitivity of the minimum angular separation to inclination errors is 1.4, similar to that for the 12-hour system. For system (ii) the angular separation changes at 0.61° perminute, giving a window of only 22 seconds for the 12 m terminal, less than that for the 12-hour orbit case. The sensitivity to inclination errors is also degraded, with a minimum separation 1.9 times the error.

It is concluded that a system based on 16-hour orbits has little to recommend it over a 3-satellite 12-hour orbit system. Significant lengthening of the transfer "window" is possible, but at the expense of double the number of active satellites, a higher launch cost per satellite and increased propagation delays.

2.7 Elliptical 24-Hour Orbits

A satellite in an inclined elliptical 24-hour orbit has the same apparent position in the sky at each apogee. The ground trace may exhibit a single cross-over, the timing of which is shown in Fig. 8 as a function of inclination for 100 km and 1000 km perigee altitudes. It canbe seen that the inclination must be at least 80° for a cross-over to occur in these cases, and that the cross-over can never occur more than 6 hours either side of apogee. The 6-hour case corresponds to 90° orbital inclination. Here again it is possible to consider a system with periodic satellite conjunctions that avoids the need for double-headed terminals. Attention has been given to two systems of this type, viz:

(i) Cross-over 4 hours from apogee, inclination 83.81°, perigee altitude 1000 km. (Three satellite system.)

(ii) Cross-over 6 hours from apogee, inclination 90°, perigee altitude 1000 km. (Two satellite system.)

The ground traces for these systems are shown in Figs. 15 and 16 respectively, together with the coverage areas. In both cases the apogee is at 22° W longitude. It is clear that both systems provide more than adequate coverage of NATO countries and of the North polar region.

The maximum duration of the circuit transfer window for a 12 m SGT is found to be 2 minutes 27 seconds for case (i) and 1 minute 42 seconds for case (ii). For a 2.4 m SGT, these times increase to about 12 minutes and 8 minutes respectively. The sensitivity of the minimum separation to inclination errors is found to be 1.6 for case (i) and 2.1 for case (ii).

A system based on 24-hour inclined orbits is attractive because of the relatively small number of satellites required and the relatively long circuit transfer windows. However, propagation delays would be twice as long as for a geostationary satellite and the launch energy required would be greater that that for either a 12-hour or a 16-hour system. The system is just as sensitive to inclination errors as the other periodic-conjunction systems considered.

3. Orbits for Linked Satellites

Only circular orbits have been considered in this category.

Here the rule is that each SGT must be able to "see" at least one satellite at adequate elevation at any time, but that different SGTs do not have to be able to see the same satellite simultaneously. n satellites equally spaced around a circular orbit define a swath of width β (measured at the centre of the earth) such that the above condition is satisfied for any station within the swath. It is found that:

$$\cos \beta/2 = \cos \theta / \cos (\pi/n) \qquad ...(14)$$

where θ is obtainable from (1).

To provide continuous coverage between latitudes lt_{min} and lt_{max} one of the following techniques can be used:

(i) n satellites equally spaced around a circular equatorial orbit such that $\beta/2 = lt_{max}$

(ii) N polar orbits (circular) each with n equi-spaced satellites such that $N*\beta = \pi$. (This gives whole-earth coverage.)

(iii) N inclined circular orbits each of n equi-spaced satellites, the inclination α being such that adjacent swaths just touch at the equator, and such that the highest latitude at which adjacent swaths overlap is lt_{max}. This will be called the "close-weave" solution. It is illustrated in Fig. 17a.

(iv) N inclined circular orbits each of n equi-spaced satellites such that the highest latitude at which adjacent swaths overlap is lt_{max} and such that the swath just reaches lt_{min} at the point on the orbit 90° from the ascending node. This will be called the "open-weave" solution. It is illustrated in Fig. 17b.

In consideration of cases (iii) and (iv), it is found that for N orbits inclined at angle α

$$lt_{max} = \gamma + \theta \qquad ...(15)$$

where:

$$\tan \gamma = \tan\alpha \cos (\pi/N) \qquad ...(16)$$
$$\tan \theta = \tan (\beta/2) / \cos \epsilon) \qquad ...(17)$$

in which

$$\sin \epsilon = \sin\alpha \sin (\pi/N) \qquad ...(18)$$

For case (iii) (close-weave) the condition for adjacent swaths to touch at the equator is:

$$\sin \alpha = \tan (\beta/2) \tan (\pi/N) \qquad \ldots(19)$$

whereas for case (iv) (open-weave) the inclination must be such that:

$$\alpha = lt_{min} + \beta/2 \qquad \ldots(20)$$

In Fig. 18, lt_{max} as derived from (15) is shown as a function of β for both open-weave and close-weave cases with N as parameter. lt_{min} has been taken as 36° in equation (20), corresponding to the most southerly latitude of the existing NATO satic SGTs. Also shown are the values of β corresponding to various values of N for the polar-orbit case.

Notice that in the close-weave case for N=2, the inclination is zero and the solution represents two superimposed equatorial swaths. Hence $lt_{max} = \beta/2$ as for the equatorial case (case(i)).

The polar orbit case for N=1 requires special consideration. To achieve whole-earth coverage β would have to be 180°, and this is not possible. However, coverage down to latitude lt_{min} can be maintained provided $\beta/2$ exceeds $\pi - lt_{min}$. Thus β_{min} = 108° for coverage down to latitude 36°.

Fig. 19 compares the total number of satellites required as functions of r_s/r_e for the four cases described above and for different values of N. It can be seen that the N=1 polar orbit case just described is optimum for $r_s/r_e > 2.6$ (10,000 km altitude), the N=2 open-weave solution is optimum between this altitude and r_s/r_e = 1.67 (4,200 km altitude) and the N=3 open-weave solution is optimum below this at least down to r_s/r_e = 1.43 (2,700 km altitude).

Examination of Fig. 18 shows that for lt_{max} = 68° (latitude of most northerly NATO SGT) the optimum solutions (minimum β) for specific values of N are as follows:

N	Solution	β_{min}
2	open weave	73°
3	open weave	[illegible]
4	open weave	38°
5	c.w. or polar	36°
[illegible]	polar	180°/N

Fig. 20 shows the number of satellites required for the optimum types of solution when r_s/r_e [illegible] 1.2. Values of [illegible] are as stated in the above table. To simplify the diagram, the discrete nature of the number of satellites has been ignored. For N > 6 the optimum type of solution uses polar orbits and in this case the minimum number of satellites (N [illegible]) can be found directly for a given r_s/r_e.

Equation (14) becomes

$$\cos \pi \; [illegible] \; (N) \cos \pi/n \; [illegible] \; \cos \; [illegible] \qquad \ldots(21)$$

In this expression the minimum value of [illegible] (N) for given [illegible] can be shown to be obtained when [illegible] (N), hence

$$n(N) \; [illegible] \; 2 \; [illegible] \; \cos^{-1} \cos \; [illegible] \qquad \ldots(22)$$

[illegible] obtained from (1) as before.

A curve derived from (22) is included in Fig. 20.

4 Conclusion

In systems where all SGTs must see the same satellite at the same time, equatorial circular orbits are not economical except in the geostationary case. Polar circular orbits are more satisfactory but the number of satellites required is still large.

Use of inclined geosynchronous orbits to extend coverage to polar regions implies at least three active satellites and double-headed terminals. Only a few links could be maintained on the same satellite all the time.

A system of three satellites in inclined, highly elliptical 12-hour orbits can provide good coverage of the NATO area and the North polar region. Propagation delays are no greater than for a geosynchronous system, but launch costs should be significantly lower then those of a system using inclined geosynchronous circular orbits.

The need for double-heading of SGTs in order to avoid outages when tranferring circuits from one satellite to another can in principle be avoided in the three-satellite, 12-hour orbit system by careful choice and maintenance of orbital inclination. However, the time available for transfer is extremely short and the allowable inclination error is very small.

A system of two active satellites in highly elliptical 12-hour polar orbits is capable of providing continuous service to the 21 existing SGTs. Double-heading would be required if outages were to be avoided. It is unlikely that the system would be able to provide continuous service to any new SGTs at latitudes below 36° N.

Systems of four or six satellites in highly elliptical 16-hour orbits have been considered, with a view to avoiding double-heading by arranging periodic satellite conjunctions. The time available for circuit transfer is increased somewhat relative to the 12-hour orbit system but at the expense of a greater number of satellites, higher launch costs and increased propagation delays (probably unacceptable for speech).

Use of 24-hour highly elliptical inclined orbits can provide continuous service in the NATO area and polar [illegible] on with as few as two active satellites. If the inclination is chosen to allow periodic conjunctions, the time available for circuit transfer is considerably longer than for either the 12-hour or the 16-hour system, though the orbital inclination needs to be just as tightly controlled. The main drawback of this system is the long propagation delay.

Regarding systems in which simultaneous visibility of the same satellite by all SGTs is not required (i.e. linked or store-and-forward systems), these have the advantage of being feasible at least in principle at arbitrarily low altitudes. Hence the launch cost per satellite can be made low. However it is found that the number of satellites required to maintain coverage of the NATO area rises very steeply as the altitude decreases. Assuming circular orbits, the minimum possible number of satellites is three but this is possible only for altitudes well in excess of geostationary height. For altitudes down to approximately 10,000 km a system of polar orbiters provides NATO coverage most economically. At this altitude the number of satellites required has risen to eight. Between 10,000 km and 1000 km altitude the "open-weave" type of orbit-system is the most economical, but by the time the altitude has fallen to 1000 km about 70 satellites are required. Below 1000 km polar orbits once again become the most economical, now giving whole-earth coverage. Large numbers of satellites are now involved. For example a system operating at 100 km altitude would require more that 1000 active satellites.

Table 1
Visibility of satellites in circular equatorial orbits

$\frac{r_s}{r_e}$	Visibility Arcs			Arc of Mutual Visibility	Min No. of Sat.
	Nor folk F3 (76W, 36N)	Bjerkvik F16 (17E,68N)	Carp F8 (76W,46N)		
2	37.8W - 114.2W	–	52.3W - 99.7W	–	–
3	23.OW - 129.OW	–	30.5W - 121.5W	–	–
4	16.5W - 135.5W	–	22.3W - 129.7W	–	–
5	12.8W - 139.2W	30.5E - 3.5E	17.6W - 134.4W	–	–
6	10.3W - 141.7W	44.3E - 10.3W	14.6W - 137.4W	–	–
7	8.6W - 143.4W	51.0E - 17.0W	12.5W - 139.5W	12.5W - 17.0W	80
8	7.3W - 144.7W	55.4E - 21.4W	11.0W - 141.0W	11.0W - 21.4W	35
9	6.3W - 145.7W	58.5E - 24.5W	9.8W - 142.2W	9.8W- 24.5W	25
10	5.5W - 146.5W	60.9E - 26.9W	8.9W - 143.1W	8.9W - 26.9W	20

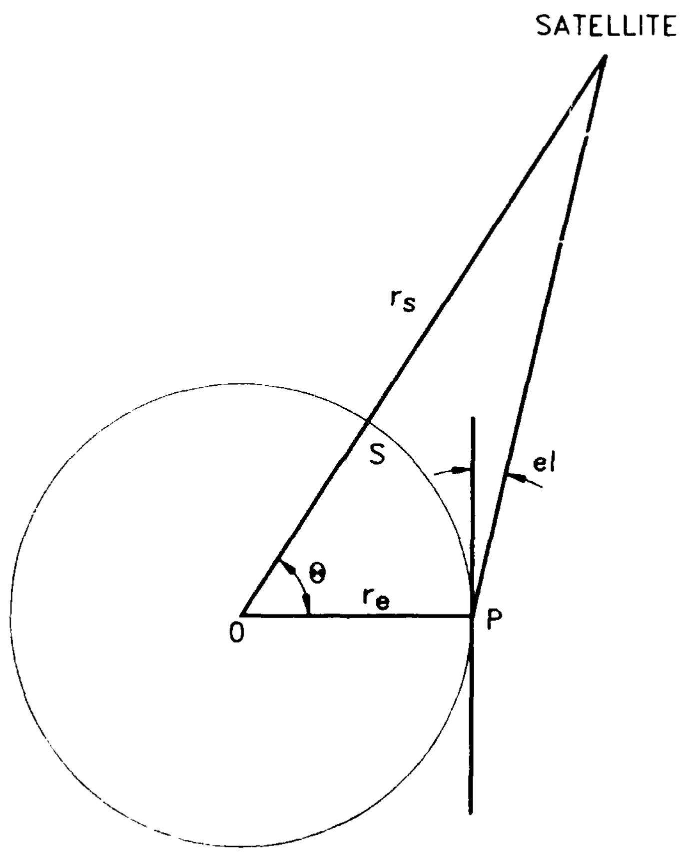


Fig 1 Basic geometry

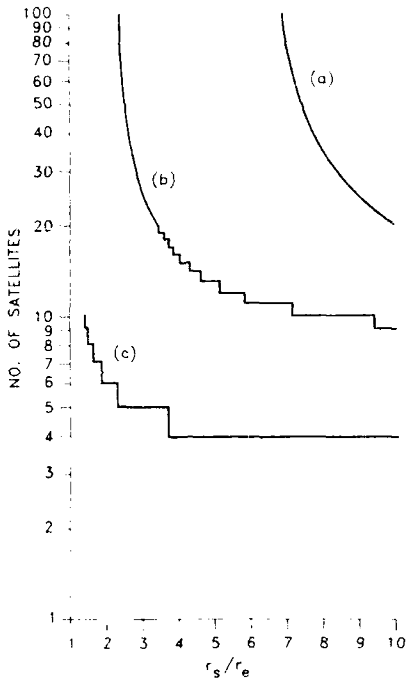


Fig.2 Number of satellites required for continuous service
(a) Using equatorial circular orbits
(b) Using circular polar orbits (for coverage above 36°N)

Fig 3 The earth seen from a geostationary satellite at 18 W showing locations of NATO SGTs

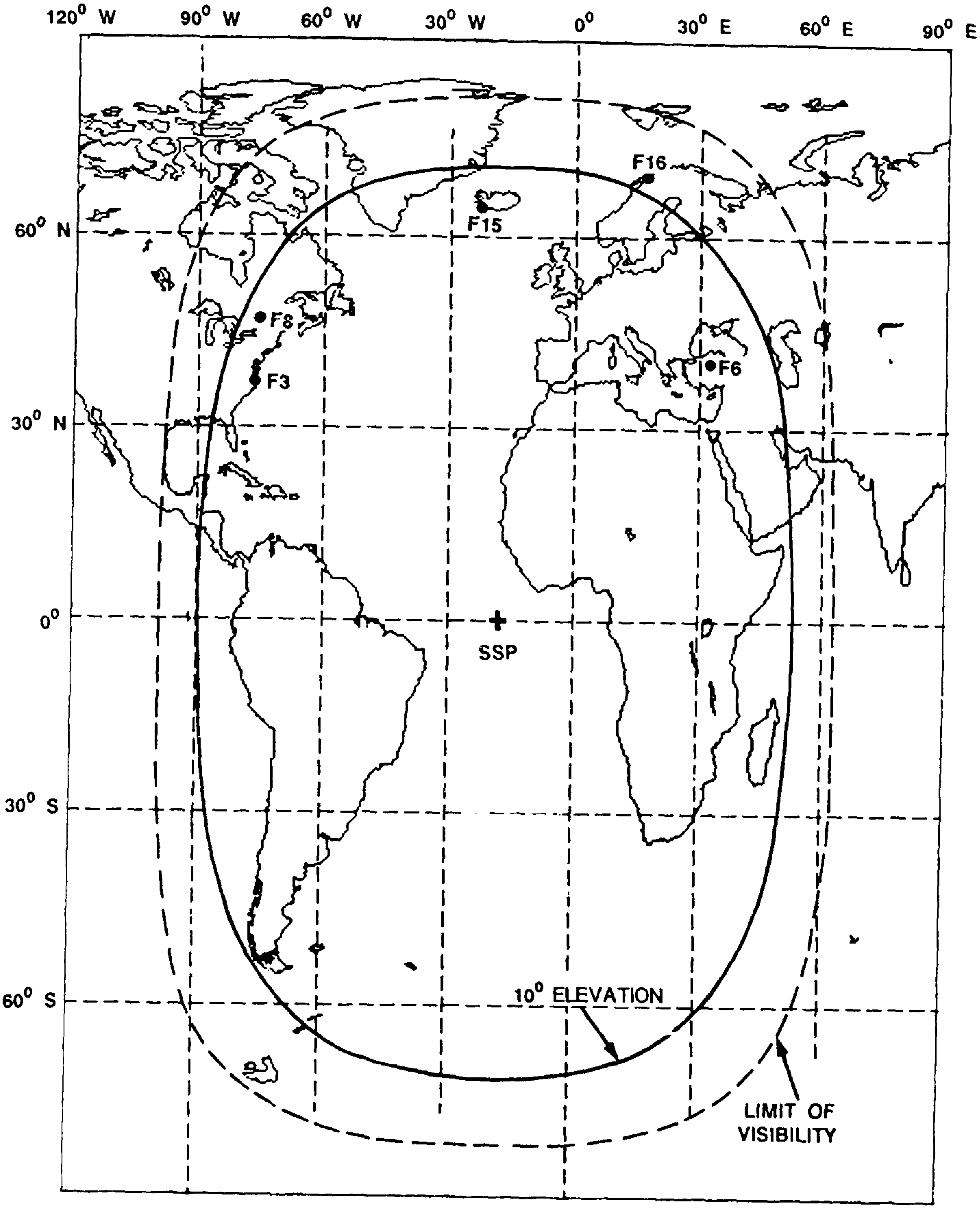


Fig 4 Coverage area of a geostationary satellite at 18°N

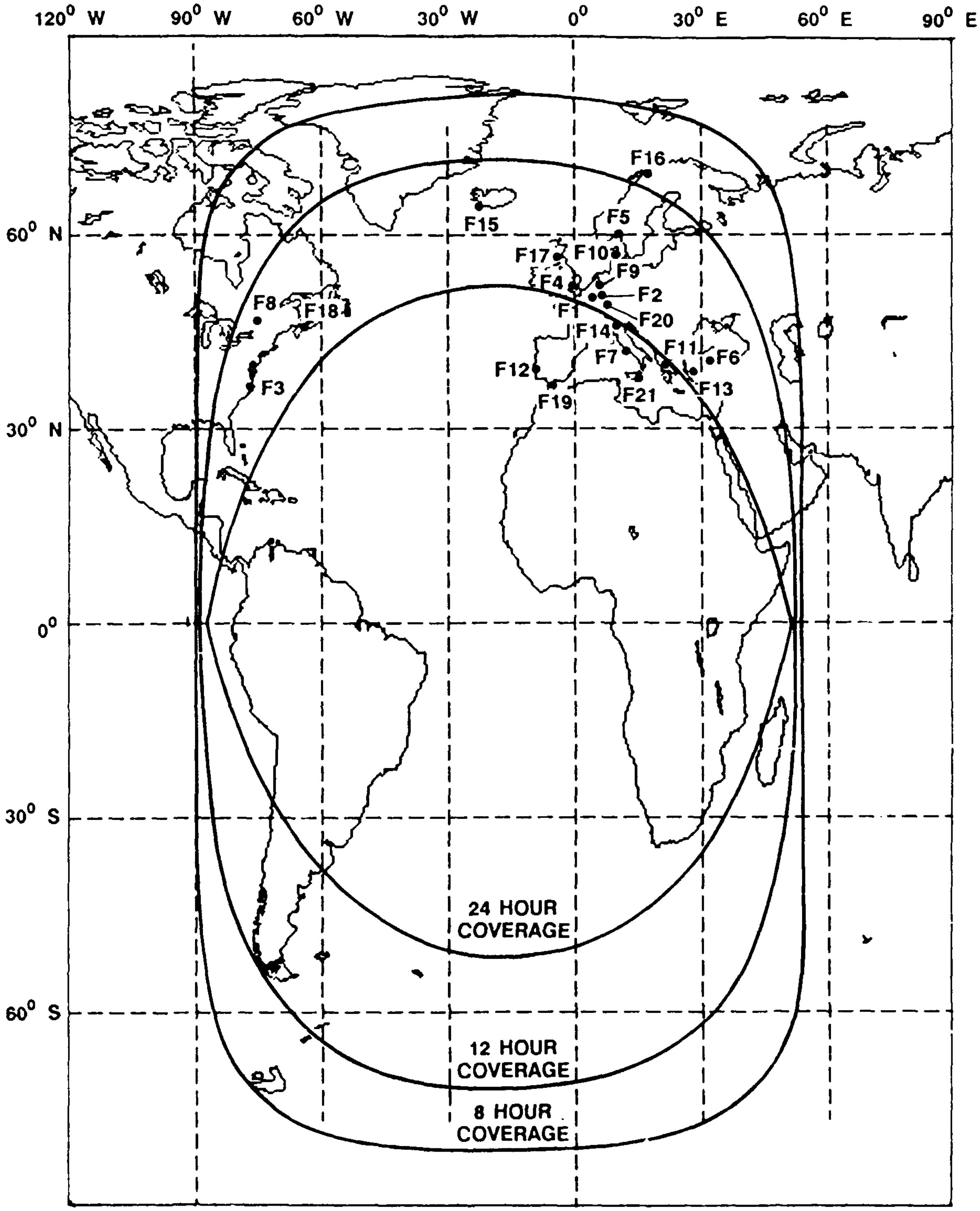


Fig 5 Coverage areas for a satellite with a 20° inclined geosynchronous orbit

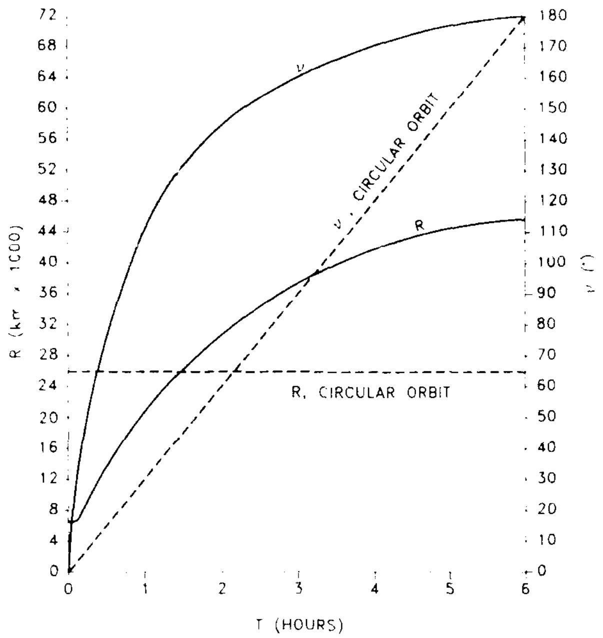


Fig 6 Time-dependence of radial vector for a satellite in a highly-elliptical 12-hour orbit (Perigee altitude 1000 km)

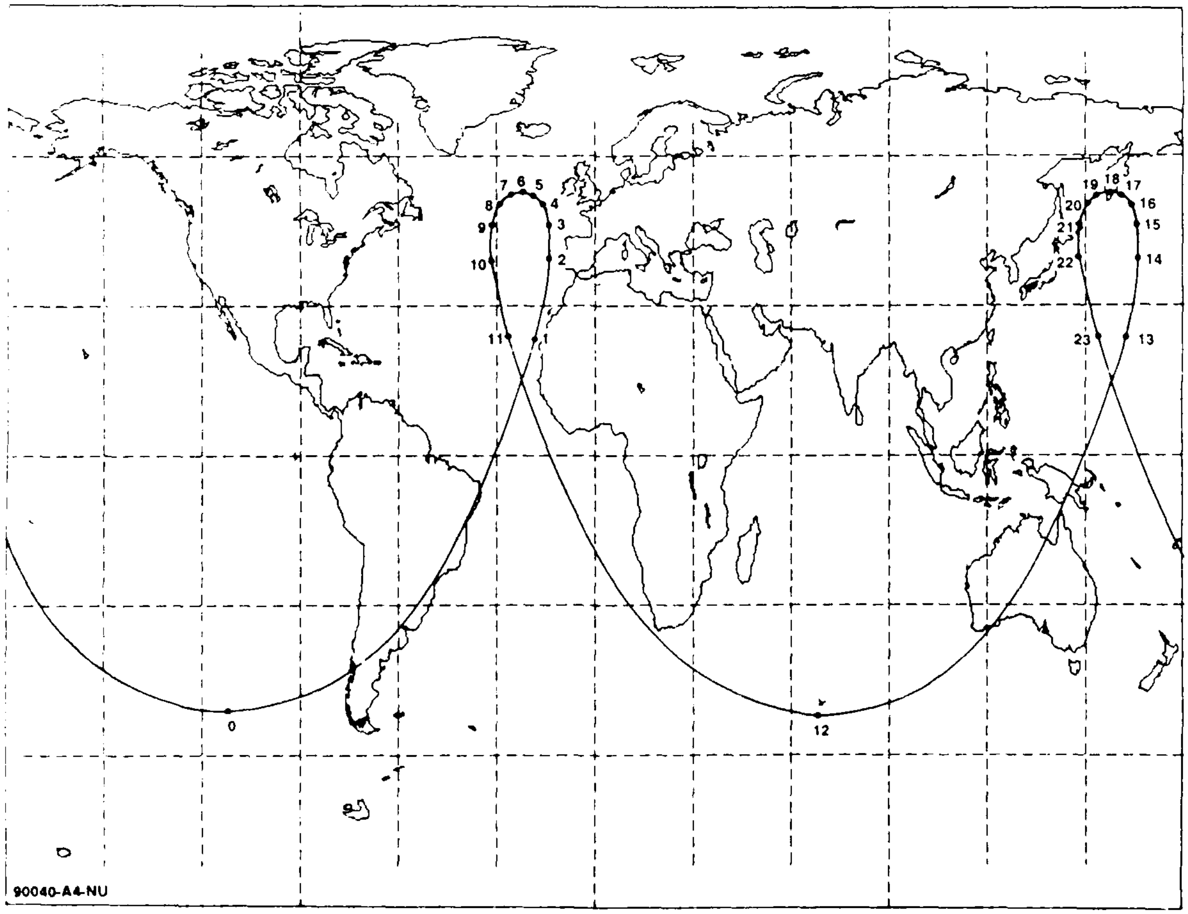


Fig.7 Ground trace of satellite in 12-hour orbit with Perigee altitude 1000 km and inclination 52°

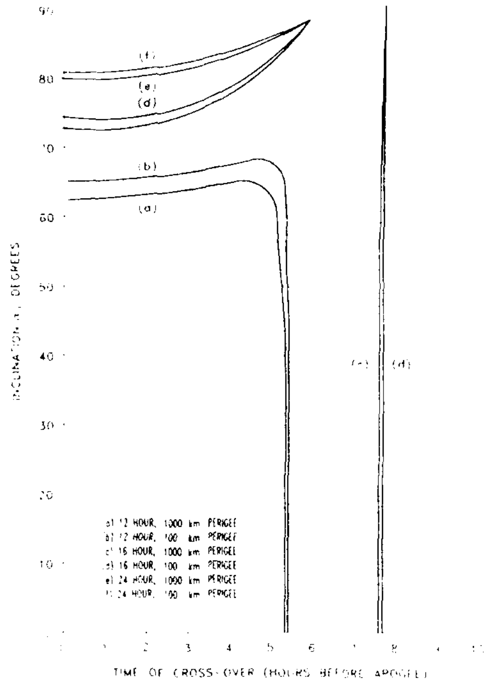


Fig.8 Timing of ground-trace cross-overs for highly elliptical orbits

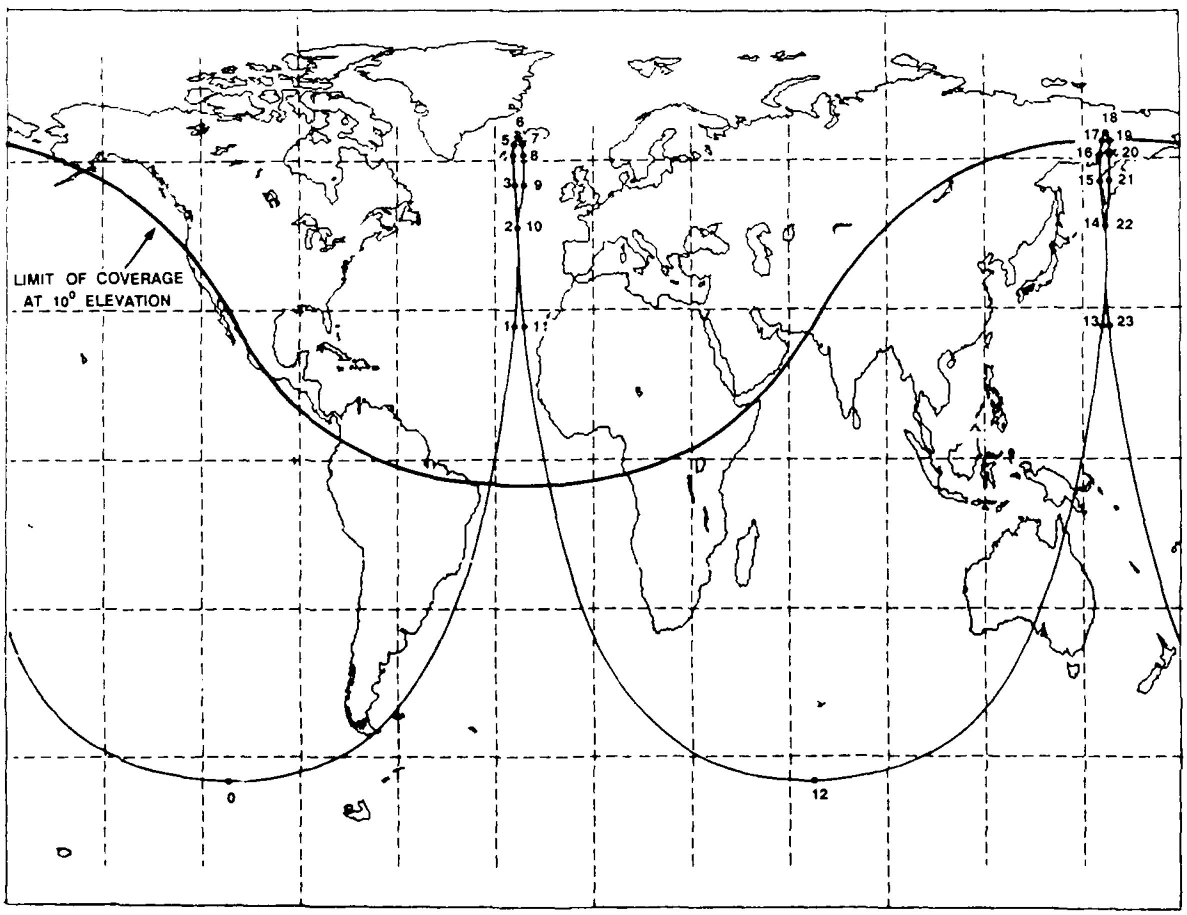


Fig 9 Ground trace of satellite in 12-hour orbit with Perigee altitude 1000 km and inclination 64.73°

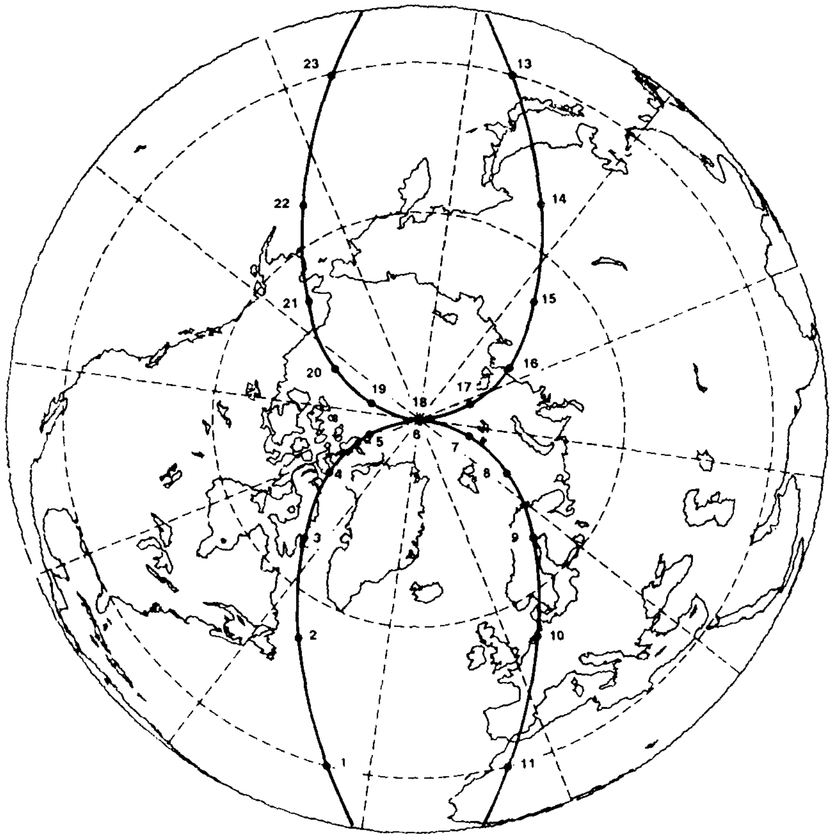


Fig 10 Ground trace of 12-hour polar elliptical orbit with 1000 km Perigee altitude

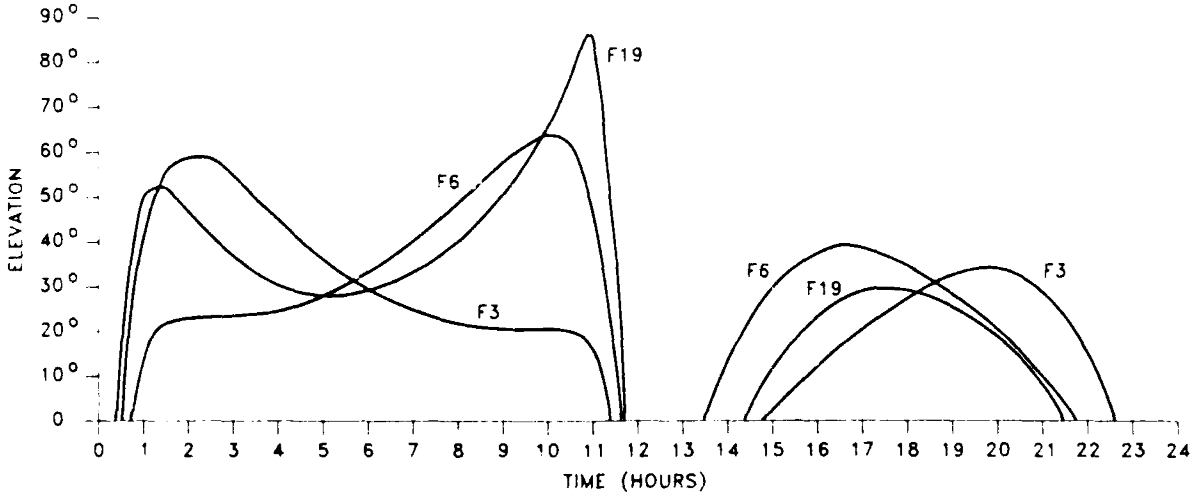


Fig.11 Elevation of satellite in polar elliptical 12-hour orbit as seen from Norfolk (F3) Ankara (F6) and Gibraltar (F19)

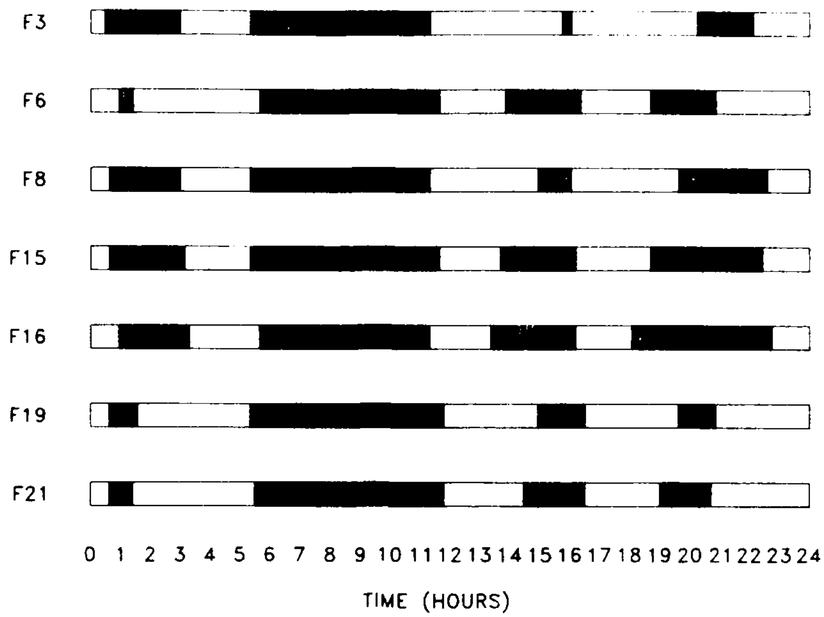


Fig.12 Periods of satellite visibility at > 10° elevation from selected SGTs for a two-satellite system using polar elliptical 12-hour orbits with 4.7 hour spacing (black bars indicate periods when both satellites are visible simultaneously)

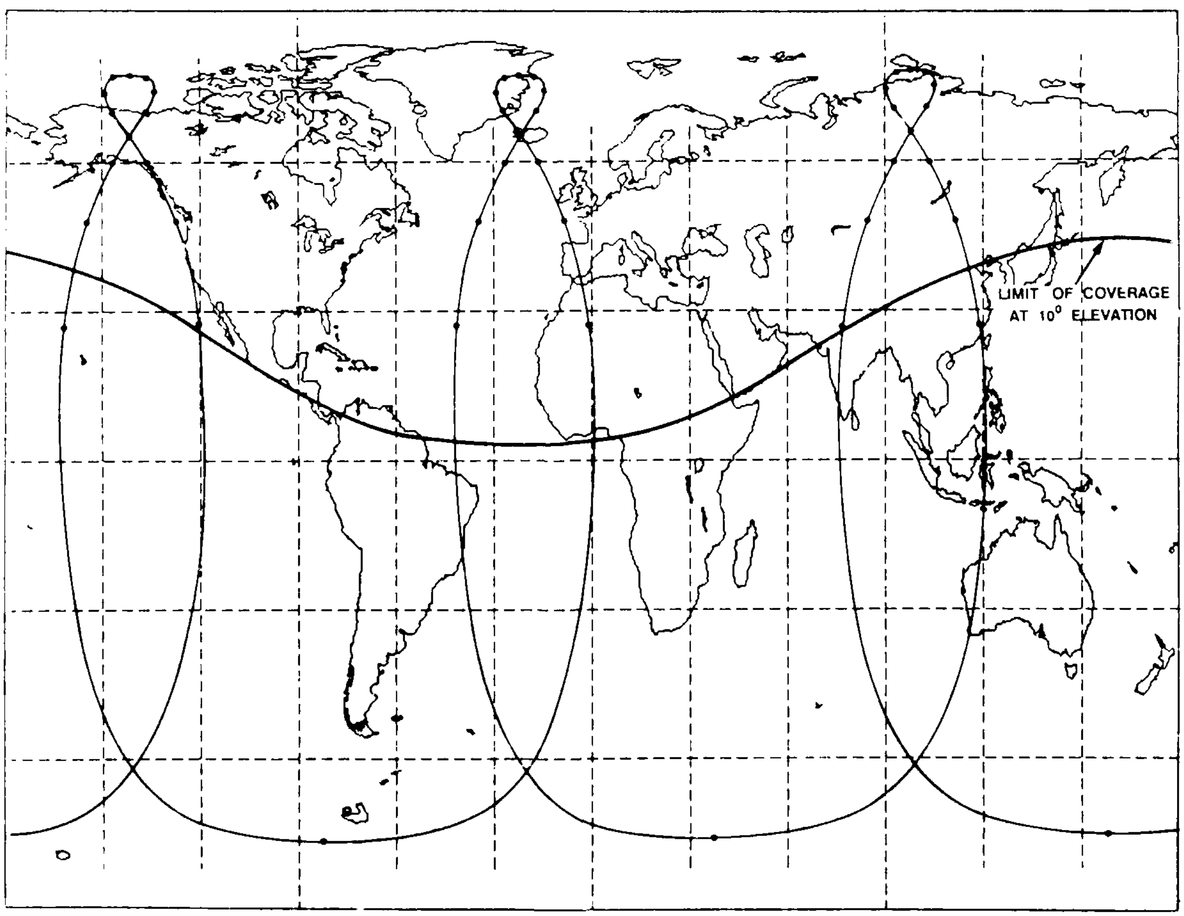


Fig 13 Ground trace of satellite in 16-hour orbit with Perigee altitude 1000 km and inclination 77.45°

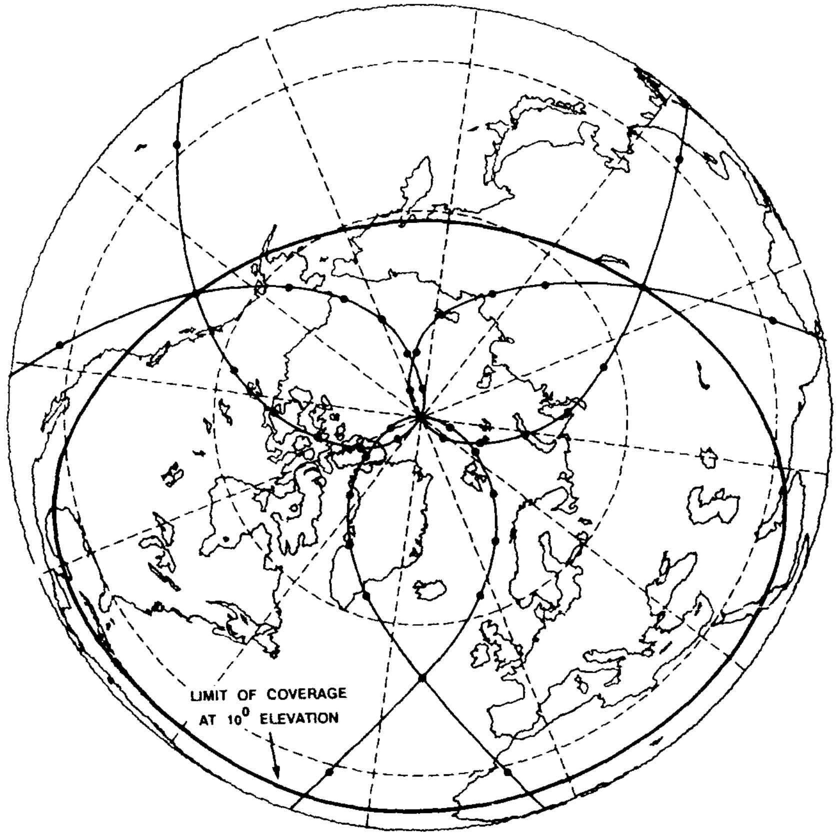


Fig.14 Ground trace of 16-hour polar elliptical orbit with 1000 km Perigee altitude

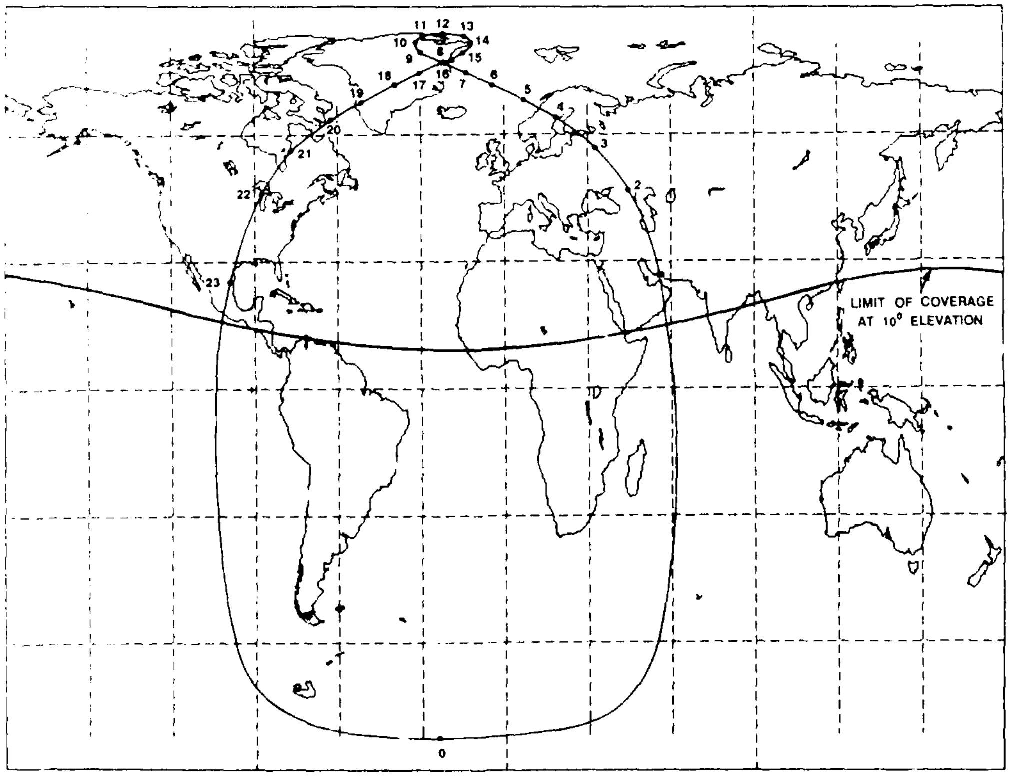


Fig.15 Ground trace of ..tellite in 24-hour orbit with Perigee altitude 1000 km and inclination 83.81°

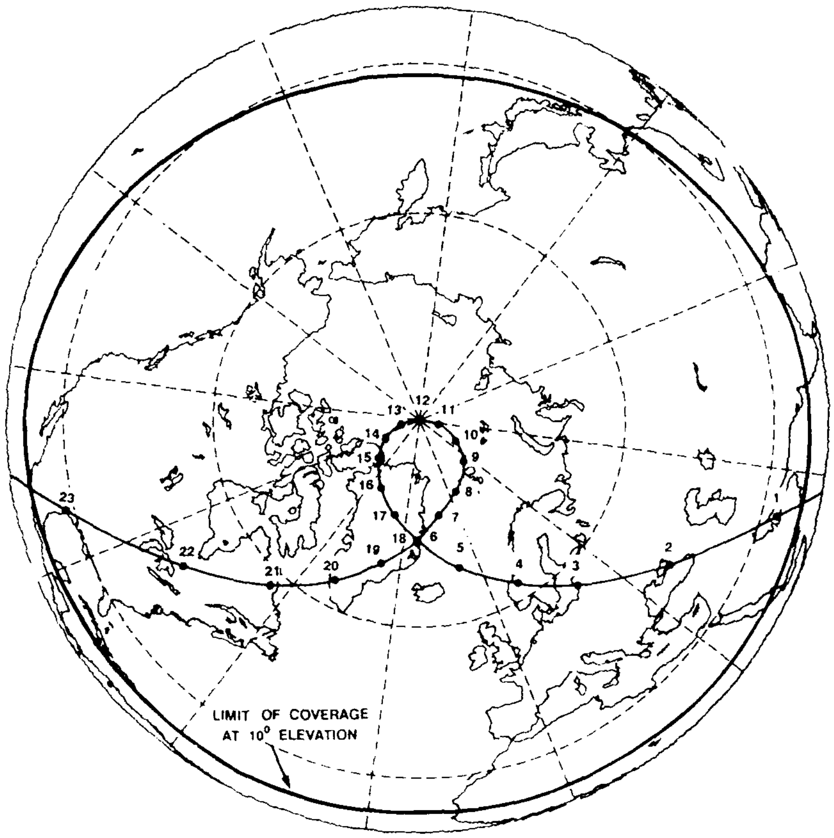
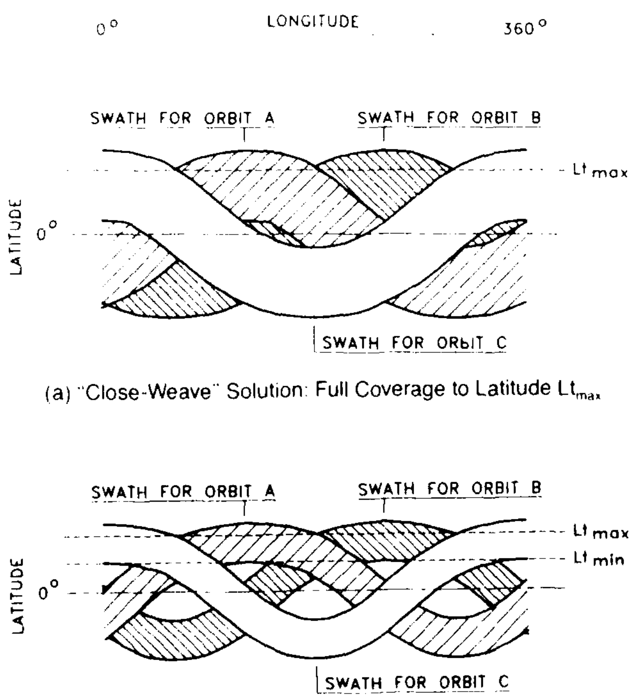


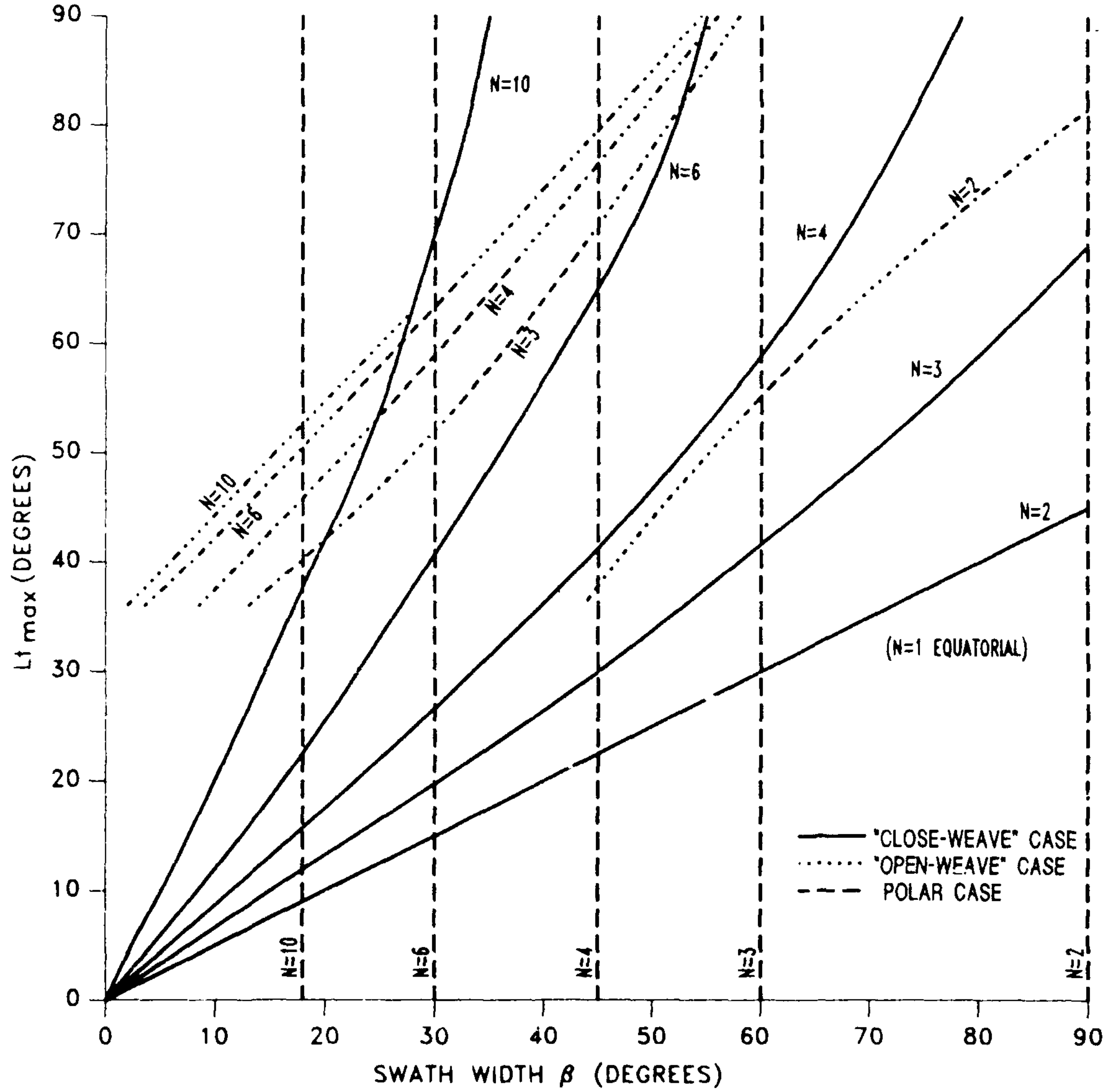
Fig 16 Ground trace of 24-hour polar elliptical orbit with 1000 km Perigee altitude



(a) "Close-Weave" Solution: Full Coverage to Latitude Lt_{max}

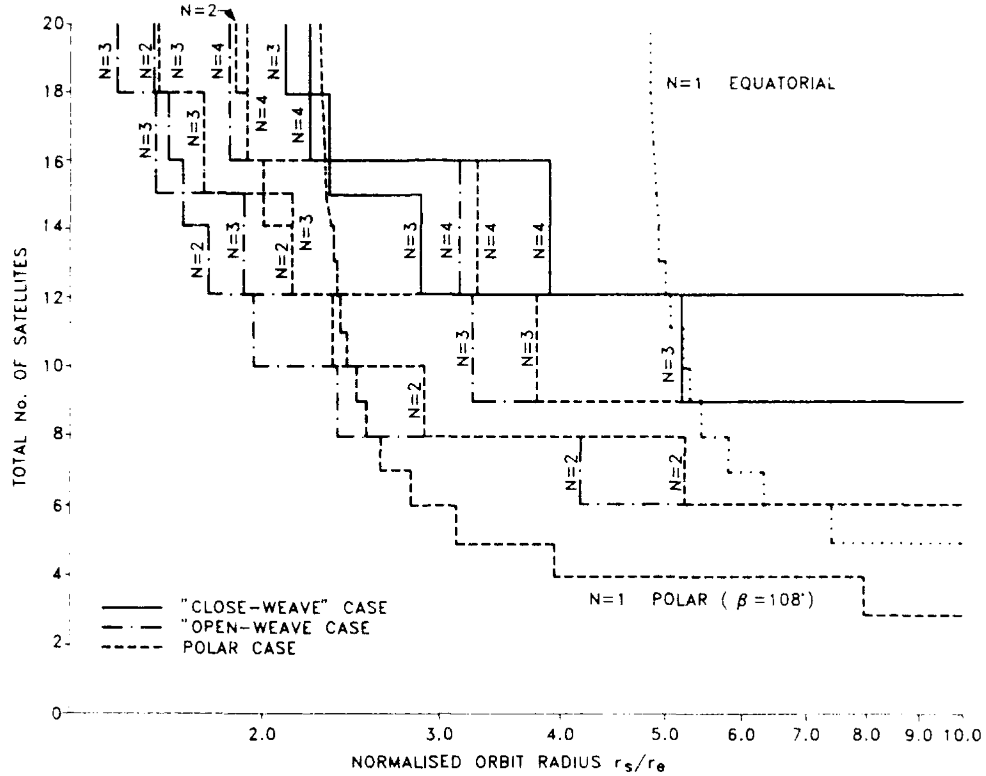
(b) "Open-Weave" Solution. Coverage between Lt_{min} and Lt_{max}

Fig 17 Coverage of the earth between given latitudes using systems of satellites in low inclined circular orbits



(—— = "close-weave" case, — · — = "open-weave" case, — — ·· = polar case)

Fig.18 Max. coverage latitude as a function of swath width



(—— = "close-weave" case, — · — = "open-weave" case, — — · — = polar case)

Fig.19 Total number of satellites as a fucntion of orbit radius for various circular-orbit systems

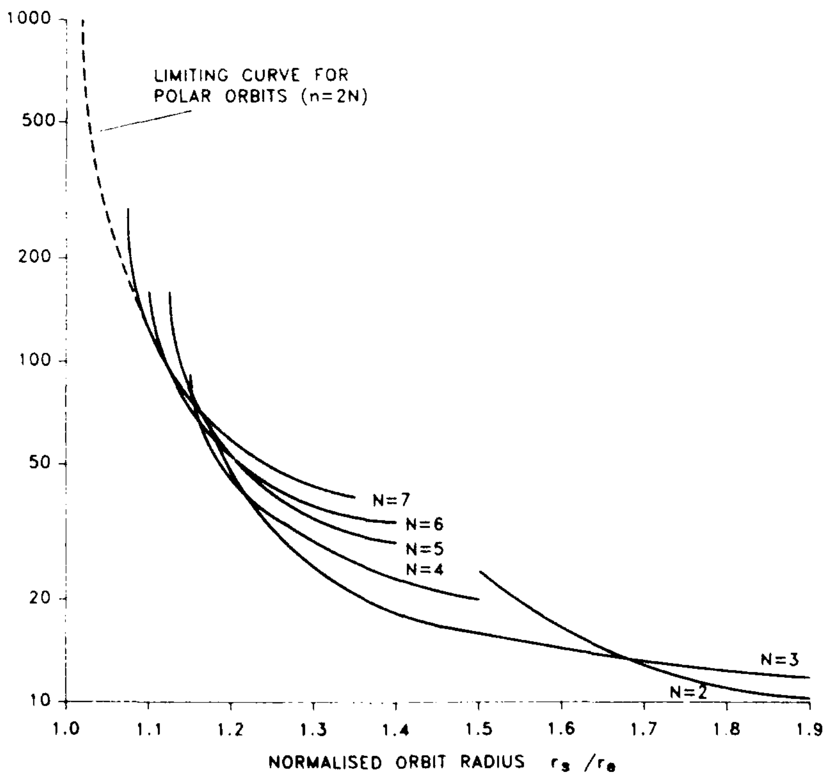


Fig 20 Total number of satellites as a function of orbit radius for optimal low-altitude circular-orbit systems

APPENDIX 10B

SOME IMPLICATIONS OF ORBITS SUGGESTED IN APPENDIX 10 A

1. BACKGROUND

Appendix 10 A describes a number of systems of orbits which might be candidates for a post-2000 NATO SATCOM system as alternatives to the quasi-geostationary orbit adopted for NATO IV and its predecessors. These options included Highly-Elliptical Earth (HEO) orbits with periods of 12, 16 or 24 hours and with various inclinations. It was shown that satellite systems using HEO orbits could give excellent coverage of the area of interest to NATO, including the North polar region. It was also shown that by careful choice of orbital parameters it is possible, at least in principle, to arrange for consecutive satellites in the same set to pass close enough to one another so that both appear, for a brief period, within the same SGT antenna beam. In this way it should be possible to maintain a continuous service without the need for "double-heading" of the SGTs. Existing satellites in HEO orbits (e.g. Soviet Molniya satellites) tend to adopt a particular orbital inclination, 63.43°, for reasons of fuel economy. This inclination is in general different from those of the HEO systems suggested for NATO in App. 10A. The aim of this Appendix is to determine the implications in terms of fuel requirements of adopting each of the proposed systems.

2. Orbit Perturbations

Due to the non-sphericity of the Earth, each of the classical parameters defining a satellite orbit changes with time. However, it can be shown that whereas the semi-major axis a, the inclination i and the eccentricity e have the same values at the beginning and end of each orbital period, the Right Ascension of the Ascending Node and the Argument of Perigee do not. Thus there are two major effects, a tendency for the orbital plane to precess about the Earth's axis and a tendency for the ellipse to rotate within the orbital plane. In the systems of interest which require the orbit to be synchronised in some sense with the rotation of the Earth, the first effect can be be compensated for by a small adjustment of the orbital period. However, the second effect will require correction, for ewample by periodic thruster firings, if the coverage and availability of the system are to be maintained.

It is shown in ref. 1 that the change in the argument of perigee $\Delta\omega$ in each orbital period is given by:

$$\Delta\omega = 3\pi J_2 (R_e/p)^2 (5\cos^2 i - 1)/2 \qquad (1)$$

where p is obtained from the relation $p = a(1-e^2)$, R_e is the mean radius of the Earth and J_2 is a constant equal to 1.082637×10^{-3}. It is evident that $\Delta\omega$ is zero when $\cos^2 i = 1/5$, i.e. when $i = 63.43°$.

3. Correcting the Change in Argument of Perigee

We need to determine the change in velocity Δv needed at the appropriate point on each orbit in order to restore the original value of ω. First, expressions are derived for components at any point on the orbit assuming it to be an umperturbed ellipse. The notation and approach is based on that of ref. 2.

With reference to Fig. 1, the satellite's position in terms of the coordinates x and y is given by:

$$x = -r\cos f = -a(\cos E - e) \qquad (2)$$

$$y = r \operatorname{isn} f = a\sqrt{(1-e^2)}\,\sin E \qquad (3)$$

where r is measured from the centre of the Earth, f is the true anomaly measured from perigee and E is refered to as the accentric anomaly. The variation of E with time is given by Kepler's equation, viz.

$$nt = E - e\sin E \qquad (4)$$

where the orbital period $P = 2\pi/n$.

Differentiating (4) w.r.t. t gives:

$$\dot{E} = n/(1-\cos E) \qquad (5)$$

Using (5), (2) and (3) may now be differentiated to give:

$$\dot{x} = \frac{a\,n\,\sin E}{1-e\cos E} = \frac{n}{u} = \frac{r\sin f}{1-e^2-e\,(r/a)\cos f} \qquad (6)$$

$$\dot{y} = \frac{a\,n\,u\cos E}{1-e\cos E} = n\,u\,\frac{r\cos f + ae}{1-e^2-e\,(r/a)\cos f} \qquad (7)$$

in which $u = \sqrt{(1-e^2)}$

With reference to Fig.2, in order to change the argument of perigee by $\Delta\omega$ we wish to impart a velocity change at the point Q where orbits 1 and 2 intersect such that the motion of the satellite at a position $\Delta\omega/2$ after apogee on the first orbit becomes that at a position $\Delta\omega/2$ before apogee on the second orbit. This is equivalent to reversing the component of velocity directed along the line of intersection OQ. Thus the change in velocity Δv is

$$\Delta v = 2\,(\dot{x}\cos\frac{\Delta\omega}{2} + \dot{y}\sin\frac{\Delta\omega}{2}) \qquad (8)$$

Substitution for $\dot{x}$ and $\dot{y}$ in (8) from (6) and (7) for the case where $f = \pi - \Delta\omega/2$ then gives:

$$\Delta v = \frac{2\,a\,n\sin(\Delta\omega/2)\,[(r/a)\;(1/u - u)\cos(\Delta\omega/2) + e\,u]}{u^2 + e\,(r/a)\cos(\Delta\omega/2)} \qquad (9)$$

For small $\Delta\omega$, $(r/a)\cos(\Delta\omega/2)$ may be approximated by $1 + e$. Expression (9) then reduces to

$$\Delta v = 2\,a\,n\,\frac{e}{\sqrt{(1-e^2)}}\sin(\Delta\omega/2) \qquad (10)$$

4. Fuel Requirements

The amount of fuel required to produce a given change in spacecraft velocity is determined by the 'specific impulse' of the fuel. This is defined as the length of time that 1 kg of fuel can produce 1 kg of thrust. If the spacecraft mass is M, the mass m_f of a fuel with specific impulse I required to produce a change in velocity Δv is therefore:

$$m_f = M\Delta v / (I\,g) \qquad (11)$$

where g is the acceleration due to gravity.

For Hydrazine used as a monopropellant (as in the NATO IV spececraft) the specific impulse is 230 seconds.

5. Implications for Specific Systems

In table 1, the change in argument of perigee per orbit the

required Δv to correct it and the equivalent mass of Hydrazine fuel assuming a 1000 kg spacecraft are stated for various HEO satellite systems. These values were determined using expressions (1), (10) and (11) respectively. The fuel required per satellite per day is also given as a basis for comparison. These estimates are undoubtedly optimistic, since additional fuel will be needed to correct second-order perturbation efects and also to maintain the correct phasing between different satellites in the same system.

It is evident that apart from the 12 - hour HEO system using 3 satellites the fuel requirements are in the range 2 - 5 kg per day. This is clearly prohibitively large when a lifetime of many years is considered. It is concluded that some alternative form of propulsion - perhaps solar sailing - would need to be used to make such schemes viable. The 0.4 kg per day requirement of the 12 - hour 3 satellite system is also impractically large even though the inclination required for periodic conjunctions of satellites - 64.73 - is close to the optimum Molniya value. It should be noted, however, that the perigee altitude assumed in this case, 1000 km, is quite arbitrary and thus it should be possible by adjusting the altitude to achieve the condition for periodic conjunctions with the optimum inclination of 63.43. This calculation has been done and the required perigee altitude is 1331 km.

6. References

[1] "Rocket Propulsion and Spacecraft Dynamics" section 18.6 (Book author and publisher unknown)

[2] "Orbital Dynamics of Space Vehicles" section 1.8 H.Deutsch, Prentice-Hall 1963

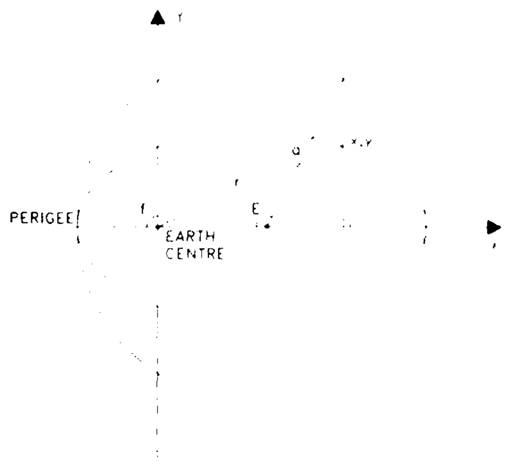


Figure 1

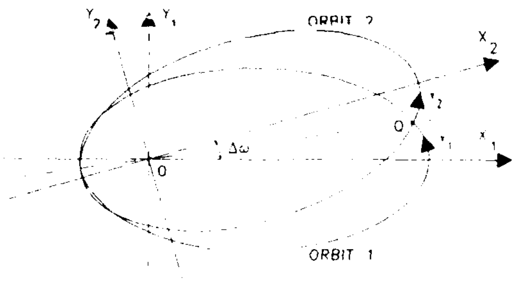


Figure 2

Table 1
Change in ω per orbit for highly-elliptical orbit systems, and fuel requirement for correction
Note: Perigee altitude is 1000 km in all cases

Period (hrs)	No. of satellites	Periodic conjunctions ?	Inc. (deg)	Δω (deg)	Δv (m/s)	Fuel per day (kg)
12	2	No	90	0.073	5.22	1.5
12	3	Yes	64.72	0.0065	0.46	0.4
16	6	Yes	77.45	0.053	3.98	2.6
16	4	Yes	90	0.069	5.18	3.4
24	3	Yes	83.81	0.061	4.81	3.1
24	2	Yes	90	0.065	5.13	2.2

CHAPTER 11

COST - PERFORMANCE ANALYSIS

11.1 SYSTEM FEATURES

So as to be able to conduct comparative cost-performance analysis of the competing architectures identified in Section 10.6 above to meet NATO's perceived requirements the following features have been assumed for the ground and space segments:

a) GROUND SEGMENT

Type	Antenna Dia (m)	Output Power (KW)	Transmission Rate(Baud)	No of simultaneous access
Shipborne	1.0	0.1	9600-2400	20
Aircraft	0.5	0.1	2400	10
Submarine	0.25	0.1	2400-100	1
Land Trans-portable	5.0 2.0	1.0 0.5	4x64000(*) 4x2400	15 30
Manpack	0.5	0.01	75	30

It should be noted that for the architectural comparison that follows the ground segment and system control are identical and do not therefore influence the cost comparison.

b) SPACE SEGMENT FEATURES

i) Nulling receive antennas
ii) Non-synchronised frequency-hopped terminals (desirable)
iii) Multibeam transmit antenna: 1 x earth cover
1 x European cover
1 x polar spot
2 x steerable
iv) On-board demodulation
v) full-bandwidth filter bank (? GHz)
vi) flexible data rate
vii) On-board re-modulation to flexible downlink format
viii) On-board routing
ix) Redefinable and assignable AJ codes
x) Simultaneous AJ multiple codes
xi) FDMA uplinks (preferred)

Figures 11.1 and 11.2 give the block diagram of two possible satellite designs which have the features outlined above.

11.2 COST-PERFORMANCE COMPARISON OF CANDIDATE ARCHITECTURES

11.2.1 Introduction

In this section comparisons are made between a limited number of candidate architectures which would meet a possible requirement that might exist in the period 2000 to 2030. The architectures evaluated are capable of being implemented to meet a wide range of possible requirements by being scaled. For simplicity, an identical requirement is assumed for each case and space segment cost for a 21 year period is used as the criteria on which each is judged. The 21 year period is implemented in three stages by spacecraft having 7 year lives, on the assumption that two changes of operational requirement, after the first would need to be implemented in the space segment. Because of the high level of investment in the ground segment a complete replacement in this sector is unlikely in less that 20 years. Cost models for architecture comparison are given in Appendix 11A.

(*) This rate can be reduced to 16 kb/s or even lower rates in the coming years without losing the quality of the 64 kb/s PCM voice

11.2.2 Candidate Architectures.

The following architectures consisting of spacecraft with the following characteristics are considered:-

(a) Dual frequency (EHF and SHF) and redundant subsystems in GEO. This architecture is used as a reference
(b) As for (a) above but operating in Inclined Elliptical orbits
(c) Separate single frequency, EHF and SHF spacecraft with redundant subsystems in GEO
(d) As for (c) above but operating in Inclined Elliptical orbits
(e) Separate single frequency, EHF and SHF spacecraft with non redundant subsystems in GEO or Inclined Elliptical orbits
(f) A mix of spacecraft in GEO and Inclined Elliptical orbits of the form in (c) and (d) above

The spacecraft types are illustrated as diagrams in Figure 11.3 below

11.2.2.1 Spacecraft Descriptions and Data

At a gross level the spacecraft can be simply divided into four main elements, an EHF payload, an SHF payload, a platform or spacecraft bus and inter-satellite links. The design of each of these is influenced by the operating orbit e.g. in transmitter power, solar array sizing, fuel capacity and inter-satellite link length. For the purposes of comparison the main parameters of interest are Mass, Power, and Reliability. Mass will be translated by simple algorithm into development and recurring cost, while spacecraft reliability will determine the number of spacecraft in the system required to meet a given theoretical availability of service.

11.2.2.2 EHF Processing Transponder Payload

a) Introduction.

The capabilities of current payloads are largely determined by dedicated hardware which implements specific processing algorithms or control functions. As a consequence it has been assumed that the mission requirements are tacitly known for the whole life time of the satellite carrying that payload. It must be expected over a period of 21 years that the requirements will change and the payload should be capable of adaptation to these needs as they arise. The design of EHF payload described will achieve this flexibility by being programmable in the areas of signal processing and switching, and antenna beam and null formation. The flexibility so achieved can be utilised by the extended life of interlinked cluster systems. The payload will embody fault tolerant designs, and have a high degree of autonomy. The transponder will utilise GaAs VLSI and VHSIC as well as MMIC all of which will be qualified by the year 2000

b) Functional Description.

EHF signals in the 44 GHz (Uplink) and 20 GHz (Down link) bands as well as on board processing using frequency hopping are used to provide robust satellite communications links ranging from 75 bps to vocoded voice 2.4 Kbps data rates to a variety of mobile terminals. For larger terminals it may be possible to support data rates as high 2 Mbps (*).

The payload provides both earth and Europe cover beams and three high gain electronically steerable transmit spot beams. Antennas with similar coverage areas are provided for reception also.

Electronically steered nulls of varying size and depth can also be positioned on jammers by programmable beam forming. The

signalling and access protocols can be made to be compatible with existing and future systems by incorporating new algorithms by programming. The upling channels (about 100) are frequency hopped Code Division Multiple Access (CDMA), m-ary frequency shift keyed (FSK) with multiple diversity. That is a symbol is transmitted over more than one frequency hop. The hop rate is adjusted to accommodate channel data rates from 75 bps to 2.4 Kbps. Higher data rates may require several symbols per hop to keep the hop rate within reasonable bounds. All channels will use error correction coding and interleaving to counter non-linear channel characteristics such as scintillation and jamming. In principle the signal formats are adaptable by virtue of its ability to reprogram the digital signal processing algorithms. It is also possible for multiple users at a large terminal to access the system by time division multiplexing on a single uplink channel. The uplink frequency hopping is removed on board, each channel remaining distinguishable by virtue of its unique hop code. Acquisition and tracking is provided on board for each channel so that access can be asynchronous. The resulting channels are demodulated, decoded and de- interleaved and passed to a baseband processor. The latter routes the resulting channel data streams to the appropriate downlink beams in response to user connectivity requests. The downlinks use a single hopped time-division multiplexed carrier for each beam. The data on these links can have a variety of modulation formats by virtue of the flexibility of the on board, modulators encoders and interleavers. These functions are carried out in the payload by a number of modules. A block diagram of the payload is given in Figure 11.4.

The implementation of each module is described in the following sections.

c) System Implementation.

i) Transmit and Receive Phased Arrays.

Separate transmit and receive arrys are required because of the large frequency difference between the transmit and receive bands (44 GHz and 20 GHz respectively). Global Europe coverage and 3 high gain spots are provided for both transmit and receive.

Each array, of diameter 0.35 m, consists of around 250 elements, each of which has an MMIC transmit or receive module in accord with the array function. The receive spot gain is 42 dBi, the transmit gain 36 dBi and the global beam 17 dBi for both receive and transmit.

Each Rx MMIC module consists of a low noise front end, followed by two stages of down conversion to an IF in the 1-3 GHz range. Each module contains amplitude and phase weighting to provide compensation for temperature and aging effects. A system noise temperature of 1500 K is predicted. Each Tx module consists of a two stage upconverter and a high power power amplifier. Again amplitude and phase weighting is provided. The first mixer receives a local oscillator signal from the down link hopping synthesizer.

ii) Digital Beam Forming, De-hopping, Acquisition and Tracking.

Both receive and transmit beam forming networks are digital, making use of stage of the art digital signal processing algorithms and VHSIC's. Incorporated with the digital beam forming network (BFN) algorithms are adaptive processing techniques which enable the efficient rejection of jammers from different locations. The output from the digital BFN will be the individually hopped channels in digital form. The dehopping acquisition and tracking may also be implemented by digital signal processing algorithm which is expected to give mass power and volume savings in the time scale considered over analogue methods. Acquisition times less than 5 seconds are predicted; allowing a user to send valid data 5 seconds after a channel has been allocated.

iii) Non-Coherent Demodulation.

The non-synchronous channels can be non-coherently demodulated by the use of FFT (Fast Fourier Transform) techniques with an anti-jamming capability. The symbol outputs from the demodulators are buffered and combined and then sent to the decoder and de-interleaver. Implementation in digital signal processing and VHSIC are envisaged. The resulting timedivision multiplexed bit stream is then passed to a digital baseband processor.

The functions described in (i), (ii), and (iii) above are carried out in a single module the Payload Processor. Changes of AJ codes, channelisation, and data rates can all be accommodated by reprogramming (See Section 8.2).

iv) Digital Baseband Processor

The digital baseband processor can be described as a time-space-time (TST) multiplexer which performs a number of basic functions on the input TDM bit stream.

- demultiplexing and buffering of the TDM bit stream
- automatic framing of data and addition of frame signalling bits (frame identity, code state, hop time, and ephemeris data)
- merging of frame as dictated by user uplink/down link connectivity requirements.
- buffering, rate conversion and routing of resulting TDM streams to the appropriate down link.

An intelligent TST multiplexer can also perform system management tasks including the following:-

- link diagnostics, i.e., traffic loads.
- link status reporting
- automatic re-configuration.

and as such it is linked to and controlled by the Payload Controller. Digital signal processing and VHSIC implementation are also used in this module.

v) Payload Controller.

This module is the central control unit for the complete payload and is essentially a high speed computer implemented in VHSIC technology. It performs the following functions:-

- telemetry and telecommand interfaces.
- common clock signals, based on a master Caesium reference to all payload modules for timing and frequency generators.
- receives ephemeris is data from the spacecraft autonomous navigation system.
- payload fault management.
- generation of code state and hop time information for each channel.
- system-wide crypto rekeying of TRANSEC keys.
- dynamically allocates and routes channels in response to user requests and traffic demands.

For extended operation with no ground intervention the autonomous navigation subsystem of the satellite and the fault management systems is essential. The navigation system can supply ephemeris data which can be multiplexed on to the down link. Ground terminals, knowing their own position (say from GPS) can then calculate their range vector to the satellite. This

(*) With reference to Section 6.1.6 (ECCM Performance of a Hypothetical EHF SATCOM System for NATO) it can be shown that the traffic requirements given in Section 11.1(a) can easily be met even under the worst uplink jamming assumed in Sec 6.1.6 with 15 dB margin.

not only gives the azimuth and elevation of the satellite for rapid spatial acquisition but also the time delay between satellite and ground terminal.

For satellites in inclined elliptical orbits, doppler information would be also be supplied for frequency tracking. The Controller also provides code state and time of hop information, for each channel which is multiplexed onto the down link for use as acquisition aids by the ground terminal

The payload fault management system is a subset of the overall spacecraft autonomous system which includes fault management and automatic re-configuration when needed. The payload fault management system contains a multilevel error and error correction system and has three main functions:-

- hardware fault detection and correction.
- software verification.
- hardware verification.

Of these hardware fault detection and correction will operate autonomously unless the condition cannot be corrected without ground intervention by telecommand. The basic mechanism to correct hardware faults will be by switching to redundant units. Software verification is a checksum operation used to verify programs in memory and is used in a continuous mode to guard against software corruption (from e.g. single event upset). Hardware verification involves a page by page check of data storage RAM. The system can have additional pages of RAM which can be mapped in place of pages showing faults. This provides a measure of software redundancy and provided the data or programs are stored in independent RAM will permit some failures in memory to exist.

vi) Local Oscillator Generators.

The generator module will used direct digital synthesis techniques using high speed GaAs circuits.

vii) Transmit Modules.

The main elements are:-

- encoding and interleaving.
- digital beam forming
- adaptive phased array.
- code generator and lownlink hopper.

The transmit array has been described in section 11.2.2.2 c(i) above the remaining modules which use digital signal processing and VHSIC are similar to their receiver conterparts. Data for an EHF redundant payload for operation in GEO is given in table 11.1.

Similar data for an EHF payload for operation in elliptical orbit (TUNDRA) is given in Table 11.2. The main difference arises from the requirement for increased transmitter power due to the higher altitude of an Elliptical orbit, e.g. TUNDRA orbit, apogee at approximately 52,000 Km as opposed to geostationary at 36,000 Km. This gives rise to a greater mass and larger raw power requirement. TUNDRA orbit has been used for cost comparison purposes since this particular orbit represents the worst case elliptical condition

The EHF payload cannot readily be split into a non-redundant form. This arises from the use of phased arrays for both transmit and receive, and the use of distributed transmit and receive modules. The failure of an individual module will have a small effect on performance and the payload will degrade gracefult rather than catastrophically.

11.2.2.3 SHF Payload

The SHF payload is assumed to be transparent and implemented in technologies similar to the EHF payload which would be available in the 2000-2030 time frame e.g. phased array antennas with distributed Tx and Rx modules and digital beam forming. The data totals for SHF Redundant payloads operating in GEO and TUNDRA orbits are given below. For the same reasons as the EHF payload an SHF cannot readily be split into a non-redundant form

For this reason in all future references to payloads of any type it should be assumed that they are redundant

11.2.2.4 Combined EHF and SHF Redundant Payload

The data for this payload operating in GEO and TUNDRA orbit is given in Table 11.3.

11.2.2.5 Inter-Satellite Links.

a) Optical link data.

The following data is based on two way optical links using semiconductor laser transmitters, MMIC and VHSIC technologies and includes:-

- pointing, acquisition and tracking equipment
- control processor unit
- interfacing to the satellite communications and attitude and orbit control systems.

The table 11.4 shows data for links from GEO to GEO GEO to TUNDRA and inter-satellite links within a cluster with up to 100 km separation-between satellites.

b) Microwave EHF Links.

EHF inter-satellite links contain similar subsystems to the optical links except that they operate in the 60 Ghz band. As a consequence the beam width of the antenna is considerably greater than that of the optical system so that the pointing and tracking requirements are less demanding and the mass and complexity of the pointing systems will be smaller. However the lower antenna gain will require more transmitter power for the same link length and this offsets some of the mass reduction. Because of the higher output power and wider beam width, the link provides a good target signal for an A-sat weapon over a wider viewing angle. The EHF system could also be easier to jam from a space based jammer. Data for EHF links are given inTable 11.5.

11.2.2.6 Data for Complete Spacecraft

Tables 11.6 to 11.12 give Mass, Power, 7-year Reliability, and Development (includes the first flight model) and Recurring Costs for spacecraft which would be used in the condidate achitectures given in Section 11.2.2. These data should not be regarded in any absolute way. They use the same algorithms for each and are thus suitable for comparative purposes.

In all cases, the spacecraft cary two Optical Inter-Satellite Links for inter-cluster communications. No other links are included. In architectures where there are no intersatellite connections the weight and cost figures given in Tables 11.6 to 11.12 should be reduced by the cost and weight figures for ISL given in those tables.

11.2.2.7 Spacecraft Launch Costs

The launch costs for each type of spacecraft is shown in Table 11.13. These are based on the most pessimistic launch costs given in Table 8.13.1 in Section 8.13 on Launch Vehicles and Space Transportation.

11.2.2.8 In Orbit Scenarios.

Applicable scenarios fall into three classes:-

(1)Geostationary (GEO).
(2)Inclined Elliptical (IE).
(3)A combination of Geostationary and Inclined Elliptical (GIE).

Each scenario may be implemented with spacecraft which are independent or inter-connected so as to share resources. The following diagrams show scenarios using spacecraft of configurations of Figure 11.3 - 1,2, and 3 only, as configurations cannot be implemented economically with payloads using phased arrays.

a) Dual frequency spacecraft in GEO.

This scenario may be implemented by one or two in orbit spare spacecraft and with or without interconnections between spacecraft. The probability of survival for 7 years and space segment costs are given below.

Two Spacecraft:-	Independent = 0.85	Cost = 415 $M	
	Interconnected = 0.93	Cost = 428 $M	
Three Spacecraft:-	Independent = 0.94	Cost = 538 $M	
	Interconnected = 0.98	Cost = 553 $M	
Replacement Costs:-	Independent = 122 $M,	Interconnected = 126 $M	

The figure also shows that changes in the EHF and SHF traffic mix during the 21 year period would be accomplished by redesign of payload and bus. This would result in recurring development costs at each stage.

b) Single-Frequency Spacecraft in GEO

This scenario may be implemented by one to two in-orbit spares of each spacecraft type. How many of each type would depend on the required mix of EHF and SHF spacecraft and availability probability. Spacecraft may also be independent or interconnected. The connectivity between spacecraft is greater for this secenario than for the dual one, as twice as many buses can be accessed for resources, leading to improved availability probabilities

An important feature of this configuration is the more gradual degratation of service which would occur in the event of a spacecraft failure and the significantly lower replacement cost of the failed craft. Availability probabilities and costs for the scenario are given below.

Four Spacecraft:-	Independent = 0.87	Cost = 489 $M	
	Interconnected = 0.93	Cost = 523 $M	
Six Spacecraft:-	Independent = 0.96	Cost = 634 $M	
	Interconnected = 0.99	Cost = 676 $M	
Replacement Costs:-	Independent = 81 $M	Interconnected = 87 $M	

The figure also shows that changes in the ratio of SHF and EHF traffic can be accomplished by replacing an SHF spacecraft in the constellation by an EHF spacecraft, thus avoiding any redevelopment of either payload or bus If necessary a single frequency satellite may use all or part of it's or it's redudant half as well as the normal operating half until an additional spacecraft of the appropriate frequency becomes available.

c) Dual Frequency Spacecraft in TUNDRA Orbit.

Operations in inclined eccentric orbits as typified by Molniya orbit are able to provide comparable continuous coverage of the NATO areas of interest to that obtained from geostationary orbit plus continuous cover of the polar areas. However a minimum of two satellites is required in a 63.4° inclined 24-hour orbit or a minimum of three satellites in a similarly inclined 12-hour orbit. To provide the same probability of service availability two or three times as many spacecraft are required as for geostationary orbit. As for the previous scenarios spacecraft may be independent or interconnected. Availability probabilities and costs for each type are given below:-

Four Spacecraft:-	Independent = 0.72	Cost = 584 $M	
	Interconnected = 0.86	Cost = 604 $M	
Six Spacecraft	Independent = 0.88	Cost = 799 $M	
	Interconnected = 0.96	Cost = 826 $M	
Replacement Cost	Independent = 92 $M,	Interconnected = 96 $M	

Costs for this scenario are at some 40 % greater than for GEO and this excess cost may be equated to the cost of improved polar area coverage.

Figure 11.7 illustrates the same method of changing EHF and SHF traffic ratios which can be used for Dual Frequency spacecraft in GEO. Redesing of the spacecraft must be carried out at each change introduction stage with consequential costs. The number of operating and in-orbit spare spacecraft may be individually altered to provide the system availability probability which is required.

d) Single Frequency Spacecraft in Inclined Elliptical Orbit.

This scenario has similar characteristics to its Geostationary counterpart but the number of spacecraft required to provide 24 hour coverage is two to three times greater. Operation in inclined elliptical orbit additionally conveys full polar coverage. Independent or interconnected spacecraft may be used. Changes in EHF and SHF traffic ratios can be implemented without spacecraft redesign and hence there are no additional recurring development costs. Service degradation resulting from spacecraft failure is more gradual and replacement costs lower than from the dual frequency scenario. Figure 11.8 shows for Stage 1 implementation a typical deploment of satellites with say 12 hour positional seperation in the orbit. The number of in-orbit spares will depend on the required availability probability. Under headings Stages 2 and 3 are shown possible variations of satellites at each orbit position to accomodate changes in traffic during the space segment life time.

Service availability probabilities and costs are given for the Stage 1 system below:-

Eight Spacecraft:-	Independend = 0.76	Costs = 692 $M	
	Interconnected = 0.87	Costs = 730 $M	
Twelve spacecraft:-	Independent = 0.93	Costs = 946 $M	
	Interconnected = 0.97	Costs = 999 $M	
Replacement Costs:-	Independent = 108 $M,	Interconnected 111 $M	

e) EHF or SHF Spacecraft in Inclined Elliptical Orbit.

In this scenario it is assumed that Polar communication coverage can be satisfactorily provided in one frequency band only. This scenario would be combined with a GEO scenario to meet the full cover requirements.

Two EHF Spacecraft:-	Independent = 0.52	Costs = 180 $M	
Four EHF Spacecraft:-	Independent = 0.85	Costs = 287 $M	
	Interconnected = 0.93	Costs = 305 $M	
Six EHF Spacecraft:-	Interconnected = 0.99	Costs = 418 $M	
Replacement Costs:-	Independent = 53 $M,	Interconnected = 57 $M	

f) Summary Data for GEO and TUNDRA Scenarios

Tables 11.14 and 11.15 collate the data given in Sections 11.2.2.8 (a) through to 11.2.2.8 (e).

11.2.3 Comparison of Architecture

11.2.3.1 General

It should be noted that the data in tables 11.14 and 11.15 is for a

7 year period. During the 21 year total period, three stages of complete space segment replacement are expected to occur, which would allow for an update for changes in traffic or other requirements. For dual frequency systems development costs will be incurred at each stage. Single frequency systems will not incur such costs since the same designs of spacecraft would be used throughout the 21 year period; only the mix of EHF and SHF types would change. Provided military components are used in the design of the spacecraft it should be possible to maintain fully availability over the 21 year interval.

An examination of the data shows that:-

a)For geostationary operations the cost of interconnection of spacecraft is of the order of 3% and is not more than 4.5% for the inclined orbits (TUNDRA as the worst case).

b)Interconnection provides significant improvement in service availability probability, ranging from 0.14 at the lower inherent spacecraft probabilities to 0.04 at the high end.

c)Increasing the system availability by the amount given in (b) above without, however, using ISL would require launching more satellites and this would increase the system cost by about 25%.

d)Operation in inclined orbits costs about 50% more than the geostationary case for the same service availability, but gives full NATO coverage including the polar region.

e)The geostationary Case 1 corresponds to the NATO IV satellites as far as coverage and the number of satellites and reliability are concerned. It is interesting to note however, that the 7-year system cost of Case 1 and that of NATO IV (about 400 $M) are almost identical even though Case 1 satellites have considerably more capacity (in SHF and EHF) and significantly greater resistance to jamming (on-board signal processing in EHF and adaptive nulling antennas).

f) The system cost changes significantly with the 7-year service availability probability. How many satellites would be needed for a 21-year period without having excessive capacity would depend on this as well as on what residual capacity would remain at the end of each seven-year period and how the change in requirements is introduced; abruptly at each 7-year period or progressively during the 21-year period. In the latter case, the calculation of the number of satellites needed for the 21-year period would require an extensive analysis of the type which was used by STC for the NATO III space segment and this would show some reduction in the total number of satellites required or, for the same number of satellites, it would indicate a lower spececraft reliability.

11.2.3.2 Comparison

Comparison data for the residual candidate architectures is given in Table 11.16

Comments on this table are as follows:-

a)The apparent lowest-cost solutions to the assumed requirements are architectures A and B. However, they use satellites in geostationary orbit and consequently do not provide polar coverage and would therefore require leased capacity if polar area cover is a firm requirement. The nature and cost of this leased capacity is unknown but would clearly add to the basic cost. It should be noted that Architecture B involves the use of inter-satellite links and an extra operating satellite to provide service availability of 0.98 which would also enhance the system resistance against jamming and physical attack.

b)Architectures C and D correspond to Architectures A and B as far as service availability is concerned but use TUNDRA orbits to provide full coverage including the polar region and consequently need twice as many satellites as for A and B costing about 50% more in all the four architectures dual-frequency payloads are used where the EHF/SHF ratio is assumed to change at 7-yearly intervals requiring R&D expenditure with the same periodicity. It is interesting to note that the average cost per payload is the lowest for Architecture D (206 $M) and highest for A (311 $M).

c)Architectures E and F are similar to A/C and B/D as far as service availabilities are concerned and they use only one type of orbit,i.e. TUNDRA, but they differ from the previous four architectures in that E and F employ single-frequency payloads (EHF and SHF). The advantage of the single-frequency payloads is that R&D costs are incurred only once in the 21-year period. It is for this reason that, even though Architectures E and F involve twice as many spacecraft as C and D respectively, their total costs are comparable.

d)The remaining architectures G,H,I,and J all use single-frequency spacecraft in TUNDRA orbits in combination with spacecraft in geostationary orbit. They, therefore, employ more satellites (9 for G and 18 for J) than in the previous architectures thus giving enhanced resistance against jamming and physical attack. They all use a mix of single-frequency spacecraft of EHF and SHF in geostationary orbit except Architecture G which uses a dual-frequency spacecraft with the disadvantages that satellite types are different for the two orbit types and spares are therefore not common and moreover the dual-frequency spacecraft need redesing at the 7 year points.

e)Architectures H and I are identical except for the inter-satellite links provided for I. It is this addition which accounts for the higher cost of I over H but in return provides a higher probability of continuity of service as well as a better AJ capability. It is assumed for both options that polar coverage provided by the use of EHF payload only is acceptable operationally. If SHF operation, in addition to EHF, is also required for coverage of the polar region then Architecture J recommends itself. This options possesses all the advantages of the other options but costs (4045 $M) much more than the others; 60% more than Architecture H which has the same service availability but costing only about 2500 $M.

11.2.4. Conclusions

A wide range of architectural options given above can be used to select the architecture which best suits the operational requirements as they will be known nearer the date of implementation of the future NATO SATCOM systems. However, based on the assumptions made regarding possible future NATO requirements, reliabilities of future electronic systems and costs per kilogram of Payloads, Spacecraft Platforms and Launches which have been used consistently for all of the candidate architectures, the following conclusions can be drawn:-

a)Provided NATO can accept the coverage provided by a geostationary only system of satellites, Architecture A is the lowest cost solution.

b)If Polar coverage obtained by leasing from the USA, costs less than Cost(H-A) = 1288 or Cost(I-A) = 1417 mil.$ then Architectures A or B plus Polar leasing would provide the next lowest cost options. Option B provides improved availability and AJ capability but at 38% higher cost than the cost of A.

c)The lowest cost architecture which provides full coverage is Architecture H at a cost increase of 50% over geostationary only.

d)For a further 5% increase in cost, an improvement from 0.95 to 0.98 in operational availability and an enhanced AJ capability can be obtained by using Architecture I. This architecture is

probably the most cost-effective option of those considered, to all of the assumed future NATO requirements.

11.3 Approaches to Cost Reduction

As was pointed out in the introduction to this report, cost will be the driving factor in determining the choice of a future SATCOM architecture. It is therefore important to consider what steps can be taken to reduce cost without changing the operational requirements.

The analysis given in the preceding section quite clearly shows that there are three predominant contributions to the overall system cost, namely:

Research and development costs
Recurring costs
Launch costs

Thus it is appropriate to consider these individually and indicate what steps could be taken to bring about cost reductions in each case.

Economy in R&D costs could be achieved through NATO/national collaboration. If NATO and national requirements can be harmonised so that a small number of spacecraft or payload types will serve either NATO or national needs then the R&D costs can be shared. A further saving to NATO would occur if two or more nations could harmonise requirements among themselves as well as with NATO so that they may share the same space segment.

Another way of achieving a saving in R&D costs would be to adopt a modular approach to system diversification and evolution. If EHF and SHF subsystems (perhaps separate types of spacecraft) are developed at the outset with a view to the mix of EHF and SHF changing as the system evolves, capacity can be replaced or enhanced in modular fashion having regard to the EHF/SHF mix required at the time without the need to develop complete new spacecraft or payloads. In the same way, if satellites in two or more types of orbit are required (say geostationary and inclined elliptical) there should be as much commonality as possible between the spacecraft designs for the different orbits. Ideally, spacecraft designs suitable for any envisaged orbit should be adopted.

Recurring costs are closely related to launch costs since both depend, as a rough approximation, on the spacecraft mass. Smaller spacecraft mean lower recurring costs but more spacecraft will of course be required. However, there will be economies of scale and thus the approach of buying more but smaller satellites is recommended. Such economies of scale will be further enhanced if the same spacecraft types are used at all stages of system evolution over a period of, say, twenty years. They will also be enhanced if the same spacecraft types are bought for national as well as NATO use.

Long-term planning is the key to achieving reductions in both R&D and recurring costs in the ways described above.

Launch costs can be minimised by reducing spacecraft mass, in particular through the exploitation of new technology. It is also important to maximise compatibility with the largest possible range of launch vehicles. This includes the possibility of utilizing "piggy-back" launch opportunities.

Interconnection of spacecraft increases system reliability and therefore tends to reduce the total number of spacecraft that need to be launched.

As a final point, it should be noted that the overall system cost figures derived in the previous section assume replacement of the space segment every seven years, regardless of whether any residual capacity remains at the end of the previous period. In practice, economies could be made by timing launches in accordance with actual changes in space segment capacity and availability, instead of in accordance with a schedule which is fixed a priori. Thus a flexible launch policy should be adopted with emphasis on the ability to launch at short notice when the need arises. This is another reason why spacecraft should be compatible with a wide range of launch vehicles.

TWO PROMISING AJ SATELLITE DESIGNS
WITH ON-BOARD PROCESSING & ROUTING

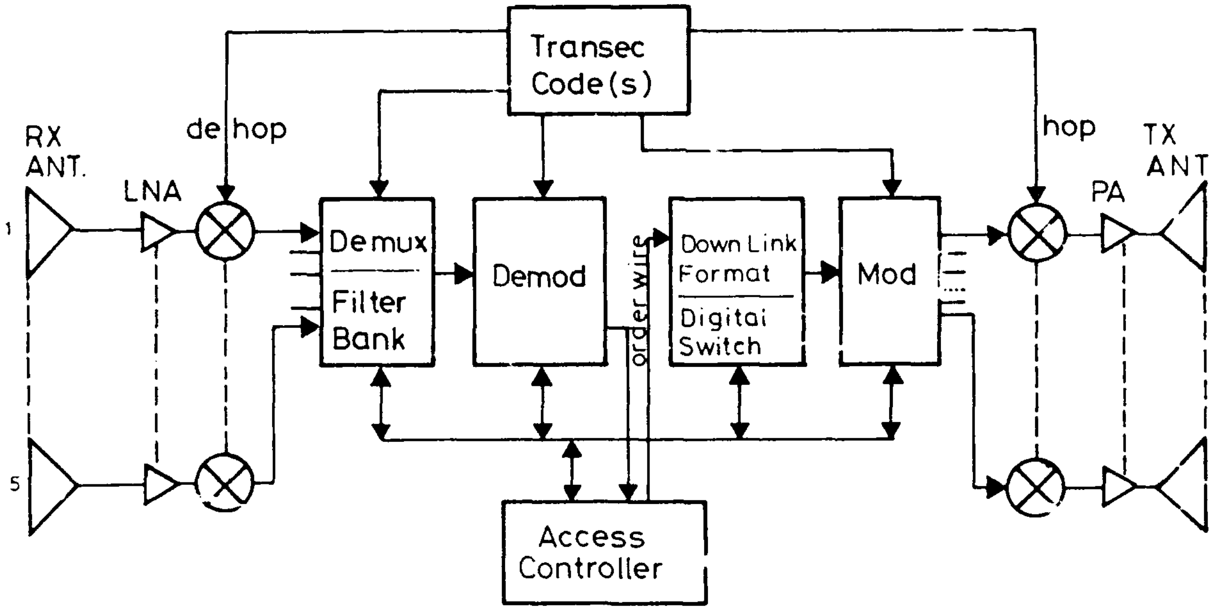


Fig. 11.1 Non-transparent satellite with about 50 MHz information bandwidth hopped across 2000 MHz. All terminals hop in synchronism. There will be as many dehoppers in the satellite as there are different nets with different codes

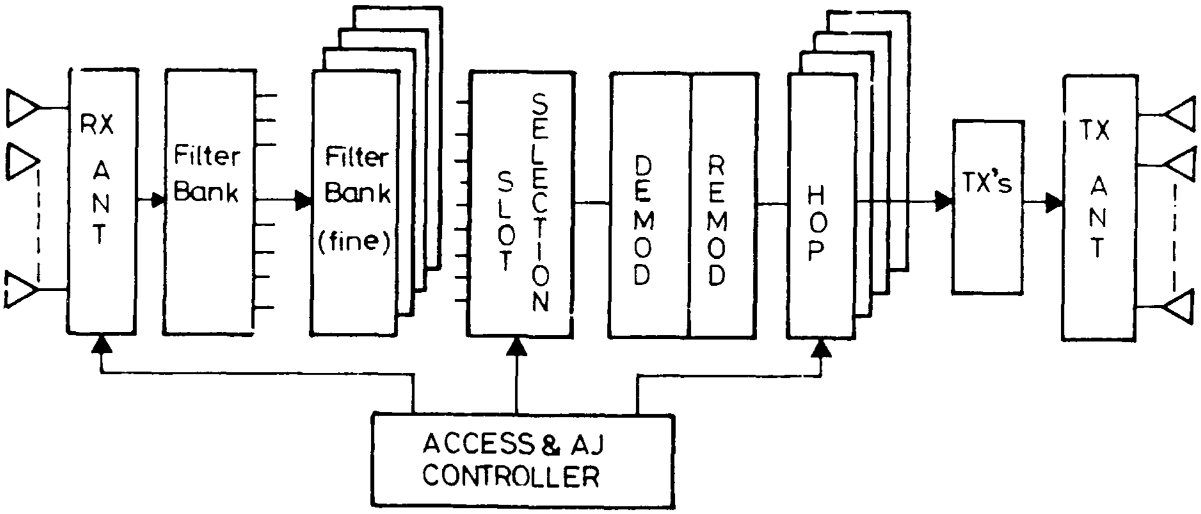


Fig. 11.2 Transparent satellite (full bandwidth) allowing operation with non-synchronised frequency-hopped terminals

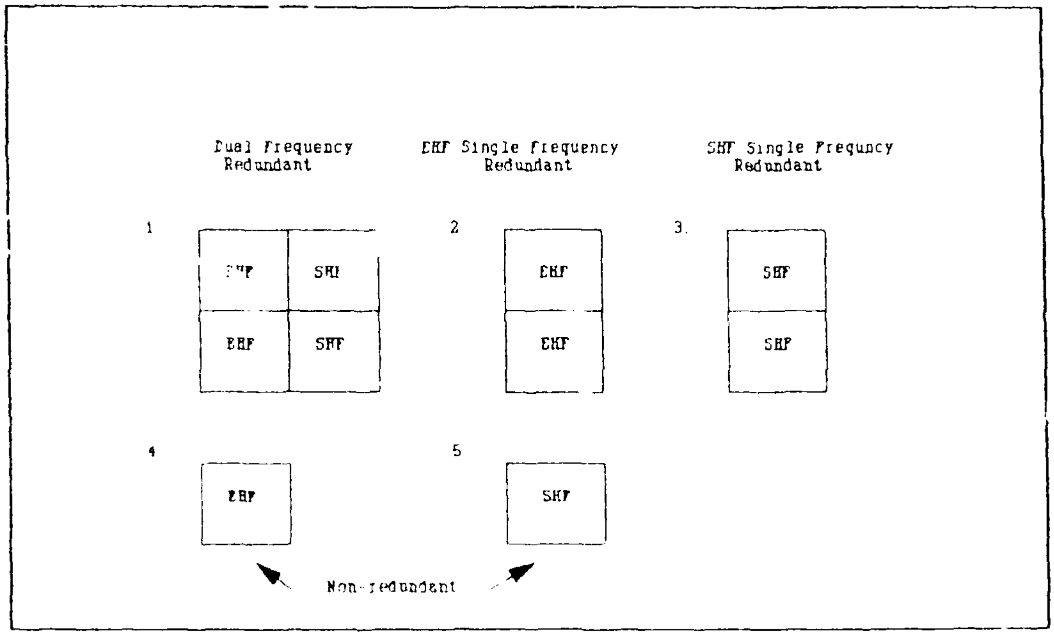


Fig 11.3 Spacecraft configurations

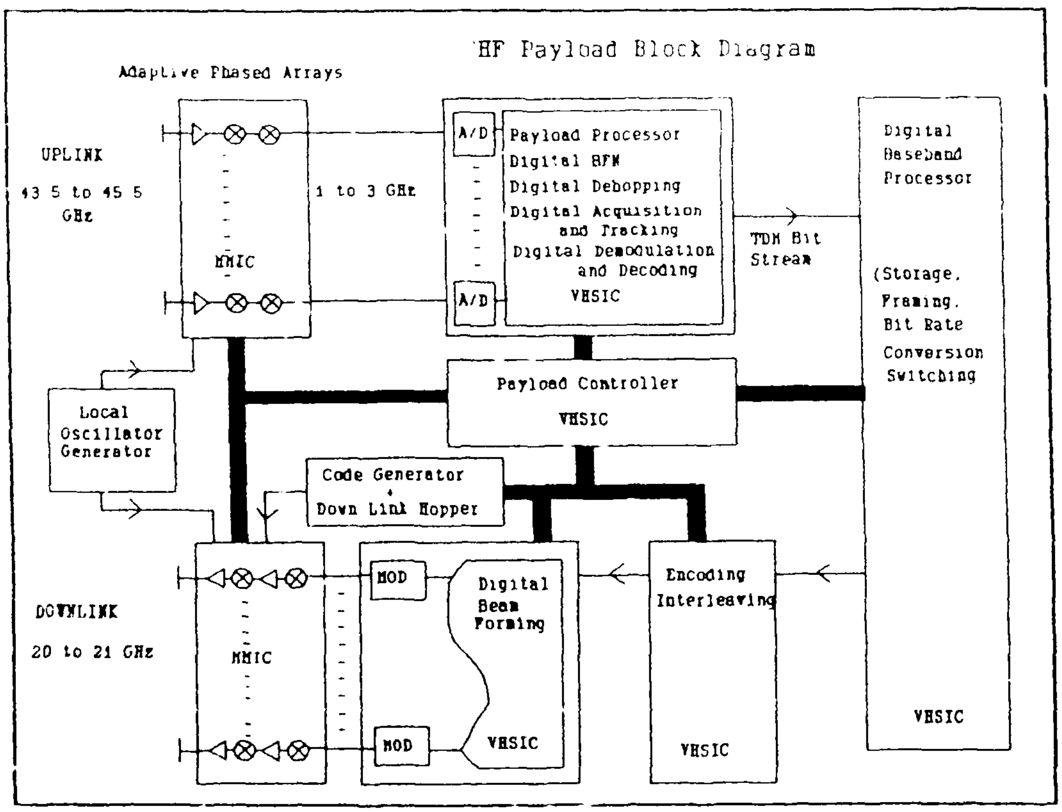


Fig 11.4 EHF payload block diagram

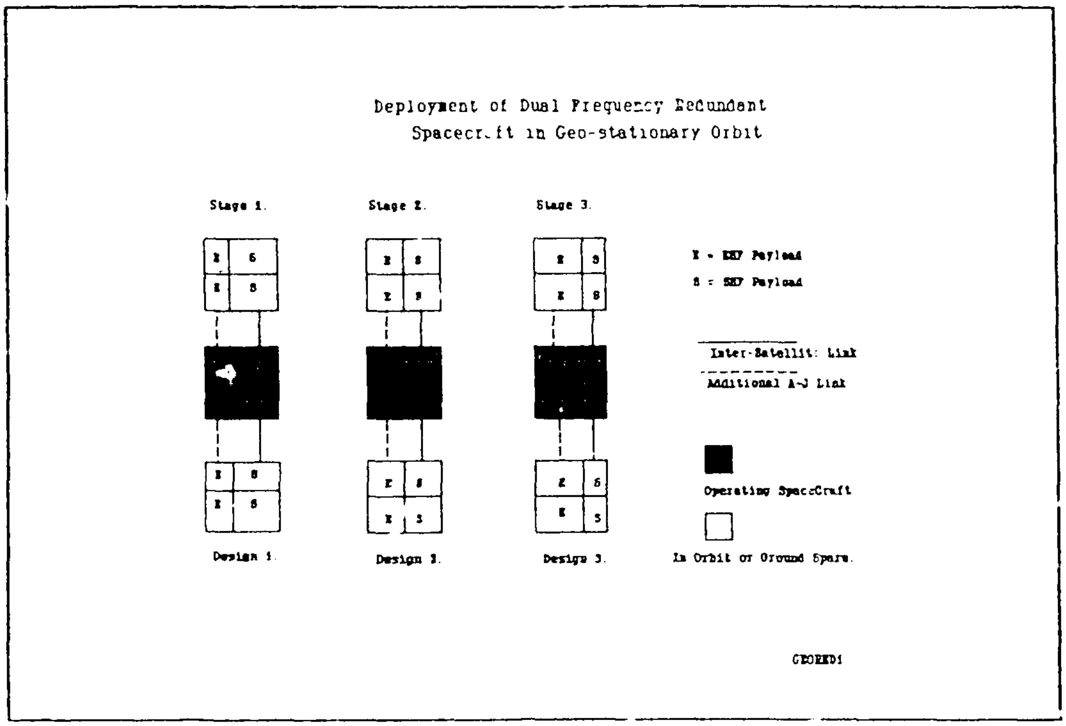


Fig 11.5 Deployment of dual frequency spacecraft in geostationary orbit

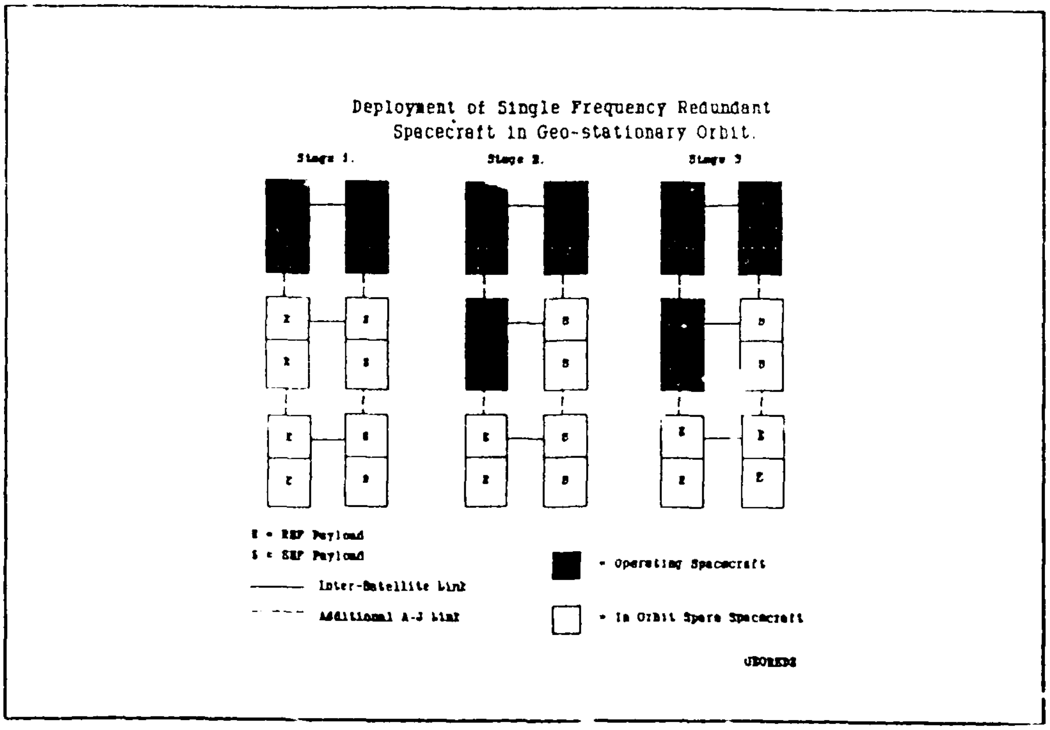


Fig 11.6 Single frequency spacecraft in GEO

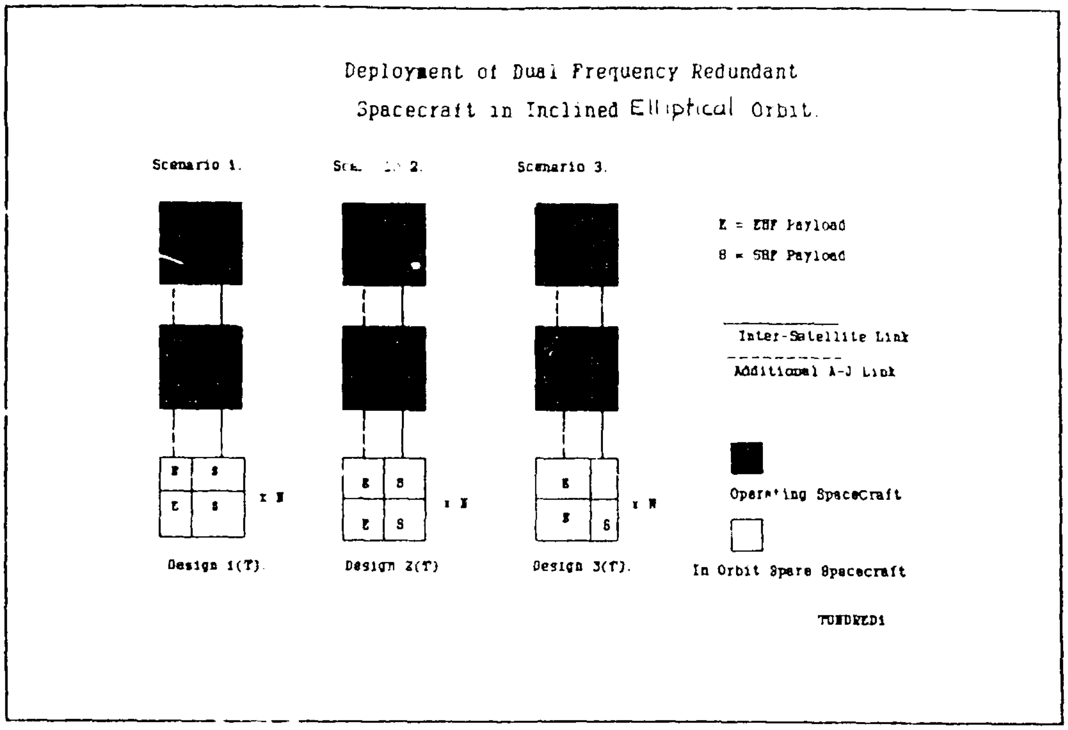


Fig 11 7 Dual frequency redundant spacecraft in GEO

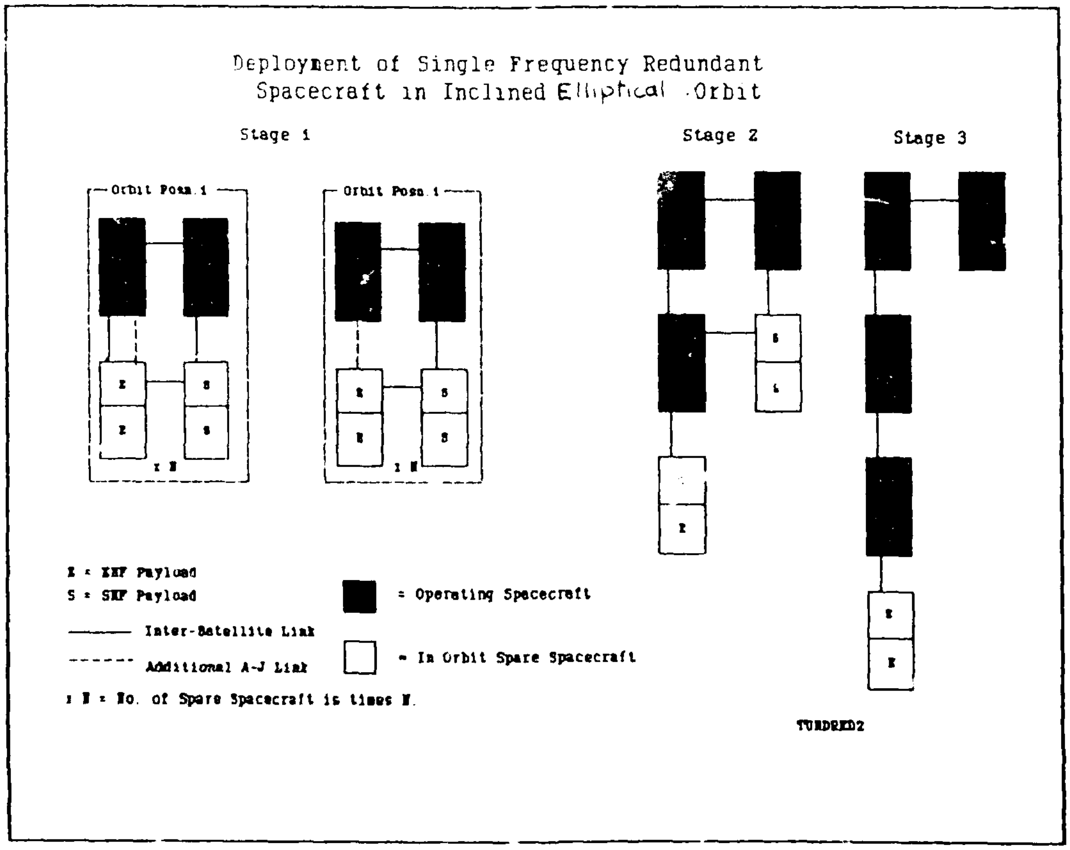


Fig 11 8 Single frequency spacecraft in inclined elliptical orbit

Table 11.1
EHF redundant payload (GEO) data

ITEM	MASS Kg.	Power Watts	Reliability 7 Years	Development * Cost $M	Recurring Cost $M
Receiver Antenna	2.5	-		0.95	0.32
Transmitter Antenna	3.5			1.33	0.44
Rx/BFR/Down Converter 250 off	25.0	156		9.5	3.20
Upconverter/Power Amplifiers 250 off	32.5	275		12.35	4.12
Local Oscillators + Ancilliary Circuits	16.0	60		6.08	2.03
Payload Processor 50 Channels	5.0	50		1.90	0.63
TOTALS	84.5	571	0.81	32.11	10.74

* Assumed development Cost/Kg 380 $M   Recurring Cost to Development Cost Ratio = 0.33

Performance Data

Global Beam Gain 17 dBi   EIRP = 16 dBW

Spot Beam Gain   G/T = 10 dBK^{-1}

Receive = 42 dBi

Transmit = 36 dBi

Table 11.2
EHF redundant payload (TUNDRA) data

ITEM	MASS Kg.	Power Watts	Reliability 7 Years	Development * Cost $M	Recurring Cost $M
Receiver Antenna	2.5	-		0.95	0.32
Transmitter Antenna	3.5	-		1.33	0.44
Rx/BFR/Down Converter 250 off	25.0	156		9.5	3.20
Upconverter/Power Amplifiers 250 off	42.5	500		16.15	5.36
Local Oscillators + Ancilliary Circuits	16.0	60		6.08	2.03
Payload Processor 50 Channels	5.0	50		1.90	0.63
TOTALS	94.5	766	0.81	35.91	12.00

* Assumed development Cost/Kg = .380 $M

Recurring Cost to Development Cost Ratio - 0.33

Performance Data

Global Beam Gain = 17 dBi

Spot Beam Gain

Receive = 42 dBi

Transmit = 36 dBi

EIRP = 18.5 dBW

G/T = 10 dBK^{-1}

Table 11.3
Combined EHF and SHF redundant payload operating in GEO and TUNDRA orbit data

Combined EHF and SHF Redundant Payload (GEO) Data.

ITEM	Mass Kg.	Power W.	Reliability 7 Years	Development * Cost $M	Recurring Cost $M
EHF Payload	84.5	571	0.81	32.11	10.74
SHF Payload	160	660	0.84	60.80	20 30
TOTAL PAYLOAD	244.5	1231	0.68	92.91	31 04

Combined EHF and SHF Redundant Payload in Tundra Orbit

ITEM	Mass Kg.	Power W.	Reliability 7 Years	Development * Cost $M	Recurring Cost $M
EHF Payload	94.5	766	0.81	35.91	12.00
SHF Payload	180.0	935	0.84	68.40	22 80
TOTAL PAYLOAD	274.5	1701	0.68	104.31	34 8

* Assumed development Cost / Kg = .380 $M   Recurring Cost to Development Cost Ratio = 0.33

Table 11 4
Optical inter — satellite link data

Link type	Link Length Maximum Km	Mass Kg	Power Watts	Reliability 7 Year	Development Cost $M *	Recurring Cost $M
Cluster Link	100	20	50	0 . 998	7 . 6	2 . 53
Inter - geostationary	73000	40	80	0 . 998	15 . 2	5 . 07
Geostationary to TUNDRA Orbit	48000	50	100	0 . 995	19 .0	6 . 33

* Assumed development Cost /Kg = 0.380 $M   Recurring Cost to Development Ratio = 0.33

Acquisition Time < 5 secs   Pointing System : - Cluster and GEO to GEO 2 Axis

Data Rate Maximum = 1 Gbps   GEO to TUNDRA 3 Axis

Table 11.5
EHF inter-satellite link data

Link Type	Link Length Maximum Km	Mass Kg	Power Watts	Reliability 7 Year	Development Cost $M *	Recurring Cost $m
Cluster Link	100	12	20	0 . 998	4 . 6	1 . 53
Inter - Geostationary	73000	26	140	0 . 998	9 . 9	3 . 3
Geostationary to TUNDRA Orbit	48000	30	238	0 . 995	11 . 4	3 . 8

* Assumed development Cost /Kg = 0.380 $M   Recurring Cost to Development Ratio = 0.33

Acquisition Time <3 secs   Pointing System : - Cluster and GEO to GEO 2 Axis

Data Rate Maximum = 1 Gbps   GEO to TUNDRA 3 Axis

Table 11.6
Spacecraft in TUNDRA orbit with EHF redundant payload

ELEMENT		MASS Kg	POWER W	Reliability 7 Year	Development Cost $M	Recurring Cost $M
PAYLOAD	SHF	0.00	0.00	-		
	EHF	94.50	7[illegible].00	0.81		
	ISL 1	20.00	50.00	0.995		
	ISL 2	20.00	50.00	0.995		
PAYLOAD SUBTOTAL		134.50	866.00	0.803	51.11	16.87
PROPULSION :		34.84	-			
AOCS :	SENSORS	12.46	34.64			
	CONTROLLER + TT&C	20.30	50.00			
POWER :	ARRAYS	65.04	25.00			
	BATTERIES	8.27	20.00			
	CONTROLLER	10.00	25.00			
THERMAL :		17.45	0.00			
STRUCTURE		61 71	0.00			
HARNESS		20.57	0.00			
MARGIN 10%		38.51	102.10			
DRY SPACECRAFT		423.65	1122.74			
PROPELLANT		272.24	-			
PLATFORM SUBTOTAL		561.39	256.74	0.90	67.37	22.23
SPACECRAFT AT LAUNCH	With ISL	695.89	1122.74	0.72	118.48	39.10
SPACECRAFT AT LAUNCH	Without ISL	655.89	1002.74	0.72	103.28	34.08

Assumed Payload Development Cost/Kg = .38 $M, Platform = .12 $M
Recurring to Development Cost Ratio = 0.33

Table 11.7
Spacecraft in geostationary orbit with EHF redundant payload

ELEMENT		MASS Kg	POWER W	Reliability 7 Year	Development Cost $M	Recurring Cost $M
PAYLOAD	SHF	0.00	0.00	-		
	EHF	84.50	571.00	0.81		
	ISL 1	20.00	50.00	0.995		
	ISL 2	20.00	50.00	0.995		
PAYLOAD SUBTOTAL		124.50	671.00	0.803	47.31	15.61
PROPULSION :		42.82	-			
AOCS :	SENSORS	12.46	34.64			
	CONTROLLER + TT&C	20.30	50.00			
POWER :	ARRAYS	50.69	25.00			
	BATTERIES	28.62	80.00			
	CONTROLLER	10.00	25.00			
THERMAL :		16.45	0.00			
STRUCTURE		62.25	0.00			
HARNESS		20.57	0.00			
MARGIN 10%		38.87	88.56			
DRY SPACECRAFT		427.83	974.20			
PROPELLANT		448.01	-			
PLATFORM SUBTOTAL		751.34	303.20	0.90	90.16	29.75
SPACECRAFT AT LAUNCH	With ISL	875.84	974.20	0.72	137.47	45.36
SPACECRAFT AT LAUNCH	Without ISL	835.84	874.84	0.72	122.27	40.35

Assumed Payload Development Cost/Kg = .38 $M, Platform = .12 $M
Recurring to Development Cost Ratio = 0.33

Table 11.8
Spacecraft in TUNDRA orbit with SHF redundant payload

ELEMENT		MASS Kg	POWER W	Reliability 7 Year	Development Cost $M	Recurring Cost $M
PAYLOAD	SHF	180.00	935.00	0.84		
	EHF	0.00	0.00	–		
	ISL 1	20.00	50.00	0.995		
	ISL 2	20.00	50.00	0.995		
PAYLOAD SUBTOTAL		220.00	1035.00	0.838	83.60	27.59
PROPULSION :		38.91	-			
AOCS :	SENSORS	12.46	34.64			
	CONTROLLER + TT&C	20.30	50.00			
POWER :	ARRAYS	74.33	25.00			
	BATTERIES	9.67	20.00			
	CONTROLLER	10.00	25.00			
THERMAL :		26.00	0.00			
STRUCTURE		81.30	0.00			
HARNESS		27.10	0.00			
MARGIN 10%		51.61	119.00			
DRY SPACECRAFT		567.68	1308.64			
PROPELLANT		367.76	-			
PLATFORM SUBTOTAL		715.34	273.64	0.90	85.84	28.32
SPACECRAFT AT LAUNCH	With ISL	935.34	1308.64	0.75	169.44	55.91
SPACECRAFT AT LAUNCH	Without ISL	895.34	1208.64	0.75	154.24	50.89

Assumed Payload Development Cost/Kg = .38 $M, Platform = .12 $M
Recurring to Development Cost Ratio = 0.33

Table 11.9
Spacecraft geostationary orbit with SHF redundant payload

ELEMENT		MASS Kg	POWER W	Reliability 7 Year	Development Cost $M	Recurring Cost $M
PAYLOAD	SHF	160.00	660.00	0.84		
	EHF	0.00	0.00	-		
	ISL 1	20.00	50.00	0.995		
	ISL 2	20.00	50.00	0.995		
PAYLOAD SUBTOTAL		200.00	760.00	0.838	76.00	25.08
PROPULSION :		48.72	-			
AOCS :	SENSORS	12.46	34.64			
	CONTROLLER + TT&C	20.30	50.00			
POWER :	ARRAYS	54.87	25.00			
	BATTERIES	31.56	80.00			
	CONTROLLER	10.00	25.00			
THERMAL :		24.00	0.00			
STRUCTURE		79.54	0.00			
HARNESS		26.51	0.00			
MARGIN 10%		50.78	88.56			
DRY SPACECRAFT		558.54	974.2			
PROPELLANT		536.55	-			
PLATFORM SUBTOTAL		895.09	236.10	0.90	107.41	35.45
SPACECRAFT AT LAUNCH	With ISL	1095.09	974.2	0.75	183.41	60.53
SPACECRAFT AT LAUNCH	Without ISL	1055.09	874.2	0.75	168.21	55.45

Assumed Payload Development Cost/Kg = .38 $M, Platform = .12 $M
Recurring to Development Cost Ratio = 0.33

Table 11.10
Spacecraft in TUNDRA orbit with EHF and SHF redundant payloads

ELEMENT		MASS Kg	POWER W	Reliability 7 Year	Development Cost $M	Recurring Cost $M
PAYLOAD	SHF	180.00	935.00	0.84		
	EHF	94.50	766.00	0.81		
	ISL 1	20.00	50.00	0.995		
	ISL 2	20.00	50.00	0.995		
PAYLOAD SUBTOTAL		314.50	1801.00	0.67	119.51	39.44
PROPULSION :		44.83	-			
AOCS :	SENSORS	12.46	34.64			
	CONTROLLER + TT&C	20.30	50.00			
POWER :	ARRAYS	116.46	25.00			
	BATTERIES	16.00	20.00			
	CONTROLLER	10.00	25.00			
THERMAL :		35.45	0.00			
STRUCTURE		109.80	0.00			
HARNESS		36.60	0.00			
MARGIN 10%		71.64	195.56			
DRY SPACECRAFT		788.04	2151.20			
PROPELLANT		513.87	-			
PLATFORM SUBTOTAL		987.41	350.20	0.90	118.49	39.10
SPACECRAFT AT LAUNCH	With ISL	1301.91	2151.20	0.61	238.00	78.54
SPACECRAFT AT LAUNCH	Without ISL	1261.91	2051.20	0.61	222.80	73.52

Assumed Payload Development Cost/Kg = .38 $M, Platform = .12 $M
Recurring to Development Cost Ratio = 0.33

Table 11.11
Spacecraft in geostationary orbit with EHF and SHF redundant payloads

ELEMENT		MASS Kg	POWER W	Reliability 7 Year	Development Cost $M	Recurring Cost $M
PAYLOAD	SHF	160.00	660.00	0.84		
	EHF	84.50	571.00	0.81		
	ISL 1	20.00	50.00	0.995		
	ISL 2	20.00	50.00	0.995		
PAYLOAD SUBTOTAL		284.50	1331.00	0.67	108.11	35.68
PROPULSION :		57.79	-			
AOCS :	SENSORS	12.46	34.64			
	CONTROLLER + TT&C	20.30	50.00			
POWER :	ARRAYS	81.71	25.00			
	BATTERIES	50.43	80.00			
	CONTROLLER	10.00	25.00			
THERMAL :		32.45	0.00			
STRUCTURE		106.14	0.00			
HARNESS		35.38	0.00			
MARGIN 10%		69.12	154.56			
DRY SPACECRAFT		760.28	1700.20			
PROPELLANT		809.90	-			
PLATFORM SUBTOTAL		1285.68	236.10	0.90	154.28	50.91
SPACECRAFT AT LAUNCH	With ISL	1570.18	1700.20	0.61	262.39	86.59
SPACECRAFT AT LAUNCH	Without ISL	1530.18	1600.20	0.61	247.19	81.57

Assumed Payload Development Cost/Kg = .38 $M, Platform = .12 $M
Recurring to Development Cost Ratio = 0.33

Table 11.12
Summary of Spacecraft data for TUNDRA and GEO orbits

SPACECRAFT TYPE	MASS Kg	POWER W	Reliability 7 Year	Development Cost $M	Recurring Cost $M
EHF Payload TUNDRA	695.89	1122.74	0.72	118.48	39.10
" " Without ISL	655.89	1002.74	"	103.28	34.08
EHF Payload GEO	875.84	974.20	0.72	137.47	45.36
" " Without ISL	835.84	874.20	"	122.27	40.35
SHF Payload TUNDRA	935.34	1308.64	0.75	169.44	55.91
" " Without ISL	895.34	1208.64	"	154.24	50.89
SHF Payload GEO	1095.09	974.20	0.75	183.41	60.53
" " Without ISL	1055.09	874.20	"	168.22	55.45
Combined EHF & SHF Payload - TUNDRA	1301.91	2151.20	0.61	238.00	78.54
" " Without ISL	1261.91	2051.20	"	222.80	73.52
Combined EHF & SHF Payload - GEO	1570.18	1700.20	0.61	262.39	86.59
" " Without ISL	1530.18	1600.20	"	247.19	81.57

Assumed Payload Development Cost/Kg = .38 $M, Platform = .12 $M
Recurring to Development Cost Ratio = 0.33

Table 11 13
Spacecraft launch costs

SPACECRAFT TYPE	MASS Kg	Launch Cost $ M
EHF Payload TUNDRA	695.89	17 40
" " Without ISL	655.89	16.39
EHF Payload GEO	875.84	21.90
" " Without ISL	835.44	20.88
SHF Payload TUNDRA	935.34	23.38
" " Without ISL	895.34	22.38
SHF Payload GEO	1044.31	26.11
" " Without ISL	1004.31	25.11
Combined EHF & SHF Payload - TUNDRA	1301.91	32.55
" " Without ISL	1261.91	31.55
Combined EHF & SHF Payload - GEO	1570.18	39.25
" " Without ISL	1530.18	38.25

Assumed Launch Cost/Kg = 0.025 $M

Table 11.14
System evaluation data for geostationary operations (7 years)

Case	Spacecraft		Payload			Availability		Total No. Operating Spacecraft	System Costs			
	No / Orbit	Frequency	EHF	SHF	Connect	S/C	Space Segment		R & D	Recur	Launch	Total
[illegible]	[illegible]	Dual	Y	[illegible]	No	[illegible]	[illegible]	[illegible]	242	82	[illegible]	[illegible]
2	2	Dual	Y	[illegible]	Yes	0.61	[illegible]	2	262	[illegible]	[illegible]	428
[illegible]	3	Dual	Y	Y	No	[illegible]	0.94	[illegible]	242	[illegible]	[illegible]	[illegible]
[illegible]	3	Dual	[illegible]	Y	Yes	[illegible]	[illegible]	[illegible]	262	173	[illegible]	[illegible]
[illegible]	[illegible]	Single	[illegible]	[illegible]	N	[illegible]	[illegible]	4	[illegible]	96	[illegible]	[illegible]
[illegible]	4	Single	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4	321	106	96	[illegible]
[illegible]	[illegible]	Single	[illegible]	[illegible]	N	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	Single	[illegible]	Y	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Table 11.15
System evaluation data for 24-hour TUNDRA orbits (7 years)

Case	Spacecraft			Payload			Availability		Total No. Operating Spacecraft	System Costs			
	No	Orbit	Frequency	EHF	SHF	Connect	S/C	Space Segment		R & D	Recur	Launch	Total
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	N	[illegible]	[illegible]	4	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4	238	236	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	N	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	N	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	N	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	N	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	N	[illegible]	[illegible]	4	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Table 11.16
Cost-reliability comparison of candidate architectures for a system lifetime of 21 years

Architecture	Cases No		Total No of operating spacecraft	No of Payload		Orbit / Freq.		Service Availability	Cost (21-year) ($M)			
	T	GEO		EHF	SHF	T	GEO		R & D	Recur	Launch	Total
A		1	2	2	2		D	0.85	3 × 242 726	3 × 82 246	3 × 76 228	1200
B		4	3	3	3		D	0.98	3 × 262 786	3 × 173 519	3 × 118 354	1659
C	2		4	4	4	D		0.86	3 × 238 714	3 × 236 708	3 × 130 390	1812
D	4		6	6	6	D		0.96	3 × 238 714	3 × 393 1179	3 × 195 585	2478
E	6		8	4	4	E,S		0.87	288	11 × 95.01 1045	2 × 40.5 489.4	1822
F	8		12	6	6	E,S		0.97	288	1615	734	2637
G	12	4	9	9	3	E	D	0.97	118 + 786 904	665 + 519 1184	313 + 354 667	2755
H	12	7	12	9	3	E	E,S	0.95	290.5	1470.2	727.1	2488
I	12	8	12	9	3	E	E,S	0.98	[illegible]	1550.9	745.3	2617
J	8	8	8	9	9	E,S	E,S	0.96	320.86	2558.2	1166.1	4045

APPENDIX 11A

COST MODELS FOR ARCHITECTURE COMPARISON

Parameters.

Let D = Development Cost.
- R = Recurring Cost
- L = Launch Cost.
- Di = Development cost of Spacecraft with ISL
- Ri = Recurring Cost of Spacecraft with ISL.
- Li = Launch Cost of Spacecraft with ISL.
- nx = no of each spacecraft type required for the 7 year mission
- m = the number of design change points in the 21 year mission life time (including the initial design)
- S_7 = The 7 year System Cost for case number c
- S_{21} = The 21 year System Cost for case letter A
- e = subscript designator for EHF single frequency spacecraft
- s = subscript designator for SHF single frequency spacecraft
- d = subscript designator for EHF and SHF dual frequency spacecraft
- C_r = Cost to replace failed in orbit spacecraft.

1. Seven Year System Cost for Geostationary Orbit Cases c. (Use Data in Tables 11.12 and 11.13)

Cases c = 2 or 4 $\quad S_7 = Di_d + (n_d - 1)Ri_d + n_d Li_d$ ...... 1.

Cases c = 1 or 3 $\quad S_7 = D_d + (n_d - 1)R_d + n_d L_d$ 2

Cases c = 6 or 8 $\quad S_7 = Di_e + Di_s + (n_e - 1)Ri_e + (n_s - 1)Ri_s + n_e L_e + n_s L_s$ 3

Cases c = 5 or 7 $\quad S_7 = D_e + D_s + (n_e - 1)R_e + (n_s - 1)R_s + n_e L_e + n_s L_s$ 4

2. Seven Year System Cost for Elliptical Orbit Cases c.

Cases c = 2 or 4. Equation 1.
Cases c = 1 or 3 Equation 2.
Cases c = 6 or 8 Equation 3
Cases c = 5 or 7 Equation 4

Cases c = 9 or 10 $\quad S_7 = D_e + (n_e - 1)R_e + n_e L_e$ ...... 5.

Cases c = 11 or 12 $\quad S_7 = Di_e + (n_e - 1)Ri_e + n_e Li_e$ ...... 6.

3. Cost of Replacement of Failed in-orbit Spacecraft.

Dual Frequency Spacecraft with ISL

$C_r = Ri_d + Li_d$ 7

Dual Frequency Spacecraft without ISL

$C_r = R_d + L_d$ 8

Single Frequency Spacecraft EHF with ISL

$C_r = Ri_e + Li_e$ 9

Single Frequency Spacecraft SHF with ISL

$C_r = Ri_s + Li_s$ 10

Single Frequency Spacecraft EHF without ISL

$C_r = R_e + L_e$ 11

Single Frequency Spacecraft SHF without ISL.

$C_r = R_s + L_s$ 12

4. Twenty One Year System Costs for Cases c.

The following cases are selected as providing service continuities which are > 0.95

<u>4.1 Inclined Orbit Cases.</u>

Case 4
$S_{21} = m[Di_d + (n_d - 1)Ri_d + n_d Li_d]$ ...... 13.

Case 8.
$S_{21} = Di_e + Di_s + [(n_e - 1)Ri_e + (n_s - 1)Ri_s] + (m - 1)[n_e Ri_e + n_s Ri_s] + m[n_e Li_e + n_s Li_s]$ 14

Case 12.
$S_{21} = Di_e + (n_e - 1)Ri_e + (m - 1)n_e Ri_e + m n_e Li_e$ ...... 15.

<u>4.2. Geostationary Orbit Cases</u>

Case 4.
$S_{21} = m[Di_d + (n_d - 1)Ri_d + n_d Li_d]$ 16

Case 7
$S_{21} = D_e + D_s + [(n_e - 1)R_e + (n_s - 1)R_s] + (m - 1)[n_e R_e + n_s R_s] + m[n_e L_e + n_s L_s]$ 17

Case 8.
$S_{21} = Di_e + Di_s + [(n_e - 1)R_e + (n_s - 1)Ri_s] + (m - 1)[n_e Ri_e + n_s Ri_s] + m[n_e Li_e + n_s Li_s]$ 18

5.0 System Cases (including Combined Orbits) for 21 Year Mission.

	USE Equation	
Case A	16	- GEO Case 1
Case B	16	- GEO Case 4
Case D	13	- Inc Case 4
Case F	14	- Inc. Case 8
Case J	$14 + 18 - (Di_e + Di_s) + (Ri_e + Ri_s)$	- Inc Case 8 + GEO Case 8
Case G	15 + 16	- Inc. Case 12 + GEO Case 4
Case H	$15 + 17 - Di_e + Ri_e$	- Inc. Case 12 + GEO Case 7
Case I	$15 + 18 - Di_e + Ri_e$	- Inc. Case 12 + GEO Case 8

Note For cases employing a mix of spacecraft with single-frequency payloads it is assumed that the two different payloads (EHF and SHF) would be developed at the beginning of the 21-year period and development cost would be incurred only once in this period.

CHAPTER 12

METHOD OF SATCOM SYSTEM ACQUISITION

The NATO SATCOM systems so far acquired have been based on national developments adapted to NATO requirements and the continuity of service (not necessarily full service) has been obtained by sharing or borrowing capacity from national systems. The national systems, in turn, have relied for continuity of service on the availability of capacity on the NATO system. Each procurement has contained an important element of R&D costs and since successive systems have been developed almost independently of each other, R&D costs have been, like the replacement cost, also recurring.

There has been a minimum of joint national R&D and use of the system and each procurement has been preceded by lengthy negotiations on production-sharing which has not, in general, satisfied, at least, some of the member countries. As a result of having independent NATO and national systems there has been considerable interoperability problems.

It is believed that the trend of system development, outlined above based on successive jumps in spending and capability with a minimum degree of general national participation should be and can be changed to meet the needs of the coming decades which may be characterized by uncertainty and shrinking military budgets requiring affordable and flexible systems.

The member countries have adequate experience within NATO and Europe and know that under these circumstances it is necessary to resort to joint R&D, procurement and use of the system while ensuring effectiveness and competitiveness for keeping the costs down.

What needs to be done jointly are:

(a) to define NATO and national requirements for satellite communications.

(b) to develop and agree on a system architecture.

(c) to delineate those technological areas which are critical and require R&D.

(d) to encourage and support companies and R&D establishments to form partnerships for undertaking R&D and production in a competitive manner.

There are good examples for collaborative R&D structures leading to production sharing and sale among NATO and European countries; EUREKA, IEPG, Airbus, ESA, Eutelsat, Intelsat etc. are but a few successful initiatives which can be taken as models for an imaginative joint-venture in satellite communications.

It is believed that the architectures evaluated and recommended in this report form a good foundation for (b) above and ensure also that the satellite designs outlined that are flexible need not change basically over a period of some twenty years or longer thus keeping the R&D and recurring costs to a minimum.

The report outlines also certain critical technologies for (c) above which need R&D. Some of these R&D topics are common to military and civilian SATCOMs; some others which are specific to military, are likely to be common to both national and NATO systems and yet another category of topics will be NATO-specific. It would therefore be necessary to make a more detailed assessment of the R&D topics and determine where R&D is a prerequisite and either can be relied upon present/future civilian developments or carried out jointly by member nations.

It is believed that the collaborative approach outlined above for SATCOM system acquisition in NATO would give tasks to all the existing bodies in NATO such as NACISA, STC, DRG, IEPG etc. and would probably not necessitate creating new structures. Cost and benefit analysis carried out in the report show that the architectures recommended can be implemented in the manner suggested above and could lead to systems which are considerably cheaper and much more effective than those we have had so far.

CHAPTER 13

CONCLUSIONS AND RECOMMENDATIONS

1. The task entrusted to WG-13 has been completed. The possible roles the SATCOM may play in the post-2000 era in NATO and national strategic and tactical communications as well as the shortcomings of the present systems have been identified. The future SATCOM systems will have to provide, as a minimum, communications for mobile forces and relaying of intelligence data in particular in an environment which can be described as uncertain. While nuclear confrontation of major powers appears less likely now than when the studies of WG-13 started in 1987, the threat posed by the proliferation of lethal weapons, local/regional conflicts and internal/external strifes by dissident groups cannot be discounted. The system architectures evaluated and being proposed in the report for further consideration by NATO have the attributes of flexibility in terms of both capacity and survivability against electronic and physical threat. The satellite design proposed is so flexible that it can, by programming, give any level of AJ protection limited by the availability of the RF bandwidth which is 2000 MHz in the EHF band. The modular design applied in the SATCOM architectures can provide also considerable flexibility in the choice of launch vehicles. All these attributes lead to a reduction of both R&D and satellite replenishment costs.

In fact, using the cost model developed in the report it is shown that the cost of acquiring and maintaining over a 21-year period a very reliable SATCOM system of the type advocated here can be almost a half of the cost of a similar system using the present-day NATO procurement method.

2. It must be pointed out with emphasis here that a SATCOM system of the type proposed in this report, if not only developed but also used jointly, would cost considerably less than the sum of the individual system costs and would give communications capability and flexibility to its users which could be much better than that they can afford on their own.

From the technology review made in Chapter 8 of the report the following areas appear as first candidates for a NATO R&D effort because they would provide solutions to the problems which are NATO-specific and can be made available within the time-frame considered for future NATO SATCOM systems:

- on-board flexible anti-jam signal processing which can be controlled by software to meet AJ requirements generally and to adapt to national modems and new modems introduced during the life-time of the satellite(s).

- adaptive nulling AJ receive antennas tailored to NATO needs and to keep costs down.

- autonomous control of the spacecraft and O&M generally using techniques of artificial intelligence, neural networks and robotics.

By sponsoring R&D activities, limited to payload technology, in the member nations even on modest scale, NATO can expect to get a better insight into and to make an impact on current technology developments carried out for civillian and military purposes.

3. Emerging technologies applicable to satellite communications have been studied and reported upon and from this, areas for significant technological thrust have been identified. It is expected that some of the R&D work needed to implement the type of systems recommended in the report would be undertaken by the nations singly or combined for national military and/or international civil SATCOM. Those R&D areas with particular interest to NATO should, therefore be carefully selected for joint action.

4. Finally, some consideration was given to the method of system acquisition and it was concluded that it would be desirable and beneficial for NATO and the nations to jointly develop, procure and use the system and that this cooperative approach could be implemented using the existing structures in NATO.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-AG-330	ISBN 92-835-0622-7	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title CONSIDERATIONS FOR NATO SATELLITE COMMUNICATIONS IN THE POST 2000 ERA

7. Presented at

8. Author(s)/Editor(s)	9. Date
Prof. Dr A. Nejat Ince	June 1991

10. Author's/Editor's Address	11. Pages
See Flyleaf	318

12. Distribution Statement This document is distributed in accordance with AGARD policies and regulations, which are outlined on the back covers of all AGARD publications.

13. Keywords/Descriptors

Avionics
Communications
Satellites
SATCOM

14. Abstract

The report is the joint work of all the members of the Group and is believed to contain a substantial amount of information, some of it original, on almost every aspect of satellite communications (system architecture, design, devices, techniques and technologies for space, ground and system control segments) which, it is hoped, will be found useful by a wide range of readers including military and political decision makers, planners, and the R&D community.

The material in this publication was assembled to report the results of the AGARD Avionics Panel Working Group 13.

AGARDograph 330
Advisory Group for Aerospace Research and Development, NATO
CONSIDERATIONS FOR NATO SATELLITE COMMUNICATIONS IN THE POST 2000 ERA
Edited by Prof. Dr A. Nejat Ince
Published June 1991
318 pages

The report is the joint work of all the members of the Group and is believed to contain a substantial amount of information, some of it original, on almost every aspect of satellite communications (system architecture, design, devices, techniques and technologies for space, ground and system control segments) which, it is hoped, will be found useful by a wide range of readers including military and political decision makers, planners, and the R&D community.

P.T.O.

AGARD-AG-330

Avionics
Communications
Satellites
SATCOM

AGARDograph 330
Advisory Group for Aerospace Research and Development, NATO
CONSIDERATIONS FOR NATO SATELLITE COMMUNICATIONS IN THE POST 2000 ERA
Edited by Prof. Dr A. Nejat Ince
Published June 1991
318 pages

The report is the joint work of all the members of the Group and is believed to contain a substantial amount of information, some of it original, on almost every aspect of satellite communications (system architecture, design, devices, techniques and technologies for space, ground and system control segments) which, it is hoped, will be found useful by a wide range of readers including military and political decision makers, planners, and the R&D community.

P.T.O.

AGARD-AG-330

Avionics
Communications
Satellites
SATCOM

AGARDograph 330
Advisory Group for Aerospace Research and Development, NATO
CONSIDERATIONS FOR NATO SATELLITE COMMUNICATIONS IN THE POST 2000 ERA
Edited by Prof. Dr A. Nejat Ince
Published June 1991
318 pages

The report is the joint work of all the members of the Group and is believed to contain a substantial amount of information, some of it original, on almost every aspect of satellite communications (system architecture, design, devices, techniques and technologies for space, ground and system control segments) which, it is hoped, will be found useful by a wide range of readers including military and political decision makers, planners, and the R&D community.

P.T.O.

AGARD-AG-330

Avionics
Communications
Satellites
SATCOM

AGARDograph 330
Advisory Group for Aerospace Research and Development, NATO
CONSIDERATIONS FOR NATO SATELLITE COMMUNICATIONS IN THE POST 2000 ERA
Edited by Prof. Dr A. Nejat Ince
Published June 1991
318 pages

The report is the joint work of all the members of the Group and is believed to contain a substantial amount of information, some of it original, on almost every aspect of satellite communications (system architecture, design, devices, techniques and technologies for space, ground and system control segments) which, it is hoped, will be found useful by a wide range of readers including military and political decision makers, planners, and the R&D community.

P.T.O.

AGARD-AG-330

Avionics
Communications
Satellites
SATCOM

The material in this publication was assembled to report the results of the AGARD Avionics Panel Working Group 13.

ISBN 92-835-0622-7

The material in this publication was assembled to report the results of the AGARD Avionics Panel Working Group 13.

ISBN 92-835-0622-7

The material in this publication was assembled to report the results of the AGARD Avionics Panel Working Group 13.

ISBN 92-835-0622-7

The material in this publication was assembled to report the results of the AGARD Avionics Panel Working Group 13.

ISBN 92-835-0622-7